DOCUMENT RESUME

ED 174 660 TM 009 536

AUTHOR Hall, Eugene R.; And Others
TITLE A Comparative Assessment of Three Methods of Collecting Training Feedback Information. TAEG Report No. 64. Final Report.
INSTITUTION Naval Training Analysis and Evaluation Group, Orlando, Fla.
PUB DATE Dec 78
NOTE 122p.; Page 17 may be marginally legible

EDRS PRICE MF01/PC05 Plus Postage.
DESCRIPTORS Competence; Cost Effectiveness; *Equivalency Tests; *Evaluation Methods; *Interviews; Job Skills; *Job Training; Military Training; *Multiple Choice Tests; Personnel Evaluation; Postsecondary Education; *Questionnaires; Rating Scales; Supervisors; Technical Occupations
IDENTIFIERS *Navy

ABSTRACT

Three methods of obtaining training feedback data from recent Atlanta Fleet technician school graduates and their fleet supervisors were compared: (1) a mailout questionnaire; (2) a structured interview; and (3) a job knowledge test. The results demonstrated that the questionnaire and structured interview procedure produced equivalent rating scale data concerning adequacy of training, frequency of task performance, and supervisors' assessments of graduate proficiency. Ratings of training adequacy and frequency of task performance obtained from school graduates were equivalent to those obtained from graduates' supervisors. In the short run, questionnaires were least expensive for data collection, and job knowledge tests were the most expensive. Over the long term, with larger populations of graduates, job knowledge tests became less expensive while structured interview costs remained high. Selection of methods for data collection must also consider the specific information needs to be met, plus the relative power of each method for producing the needed information. (Author/MH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

TAEG

TRAINING ANALYSIS AND EVALUATION GROUP

TAEG REPORT NO. 64

A COMPARATIVE ASSESSMENT OF THREE METHODS OF COLLECTING TRAINING FEEDBACK INFORMATION

ED174660

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

FOCUS ON THE TRAINED MAN

TM009 536

ED FOR PUBLIC RELEASE;
ITION IS UNLIMITED.

DECEMBER 1978

TRAINING ANALYSIS AND EVALUATION GROUP
ORLANDO, FLORIDA 32813

TAEG Report No. 64

A COMPARATIVE ASSESSMENT OF THREE METHODS OF COLLECTING TRAINING FEEDBACK INFORMATION

Eugene R. Hall
Carol F. Denton
M. Michael Zajkowski

Training Analysis and Evaluation Group

December 1978

GOVERNMENT RIGHTS IN DATA STATEMENT

Reproduction of this publication in whole or in part is permitted for any purpose of the United States Government.

Alfred F. Smode

ALFRED F. SMODE, Ph.D., Director,
Training Analysis and Evaluation Group

Worth Scanland

WORTH SCANLAND, Ph.D.
Assistant Chief of Staff for
Research and Program Development
Chief of Naval Education and Training

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER TAEG Report No. 64	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A COMPARATIVE ASSESSMENT OF THREE METHODS OF COLLECTING TRAINING FEEDBACK INFORMATION		5. TYPE OF REPORT & PERIOD COVERED Final Report 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Eugene R. Hall, Carol F. Denton, and M. Michael Zajkowski		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Training Analysis and Evaluation Group Orlando, FL 32813		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE December 1978 13. NUMBER OF PAGES 116
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Training
Training feedback
Training evaluation
Evaluation
Questionnaires
Job knowledge tests
Structured interviews
Methods of evaluation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A study was conducted to compare three methods of obtaining training feedback data from the fleet: a mailout questionnaire, a structured interview procedure, and a job knowledge test. The results of the comparisons demonstrated that the questionnaire and structured interview procedure produced equivalent rating scale data concerning adequacy of "A" school training, frequency of task performance, and supervisors assessments of graduate proficiency. Ratings of training adequacy and frequency of task

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT (continued)

performance obtained from school graduates were equivalent to those obtained from the graduates' supervisors.

In the short run, questionnaires are least expensive for data collection, and job knowledge tests are the most expensive. Over the long term, with larger populations of graduates, job knowledge tests become less expensive while structured interview costs remain high. However, selection of methods for data collection must also consider specific information needs to be met plus the relative power of the methods for producing the needed evaluation information.

ACKNOWLEDGMENTS

Collection of data for the study reported herein required the support of many Naval activities. In this regard, the cooperation and assistance of the staffs of the Commander in Chief, U.S. Atlantic Fleet; Commander, Naval Surface Force, U.S. Atlantic Fleet; and Commander, Training Command, U.S. Atlantic Fleet are gratefully acknowledged. The assistance of the Central Test Site for the Personnel and Training Evaluation Program in providing information and consultation concerning details of job knowledge test development and use is also acknowledged.

TABLE OF CONTENTS

Section		Page
I	INTRODUCTION.	11
	Background.	11
	Purpose	12
	Organization of the Report.	12
II	TECHNICAL APPROACH.	13
	Study Rationale and Limitations	13
	Methodological Design	13
	PEP Study Plan	14
	AE Plan	14
	BT Plan	15
	Description of the Instruments	15
	Q and SI Data Collection Instruments.	16
	JKT	16
	Correspondence of Instrument Items	19
	Instrument Development	19
	Q/SI Instruments.	19
	JKT Development	20
	AE JKT	20
	BT JKT	20
	Administration of Instruments.	21
	Questionnaires.	22
	SI/JKT General Procedures	22
	SIs.	23
	JKTs	23
	Data Handling/Processing	24
	Q/SI Forms.	24
	JKTs.	25
	Correlational Analyses	25

TABLE OF CONTENTS (continued)

Section		Page
	Resource Accounting	28
	Administrative and Technical Problems	29
III	RESULTS AND DISCUSSION OF DATA EQUIVALENCE ANALYSES	31
	Results	31
	AE Q and SI Interscale Correlations	31
	AE JKT and SI Correlations	31
	AE Subarea Correlations	33
	BT Q and SI Interscale Correlations	34
	BT JKT and SI Correlations	34
	BT Subarea Correlations	36
	Interpretation and Discussion of Results	37
	Q versus SI Method (AE)	37
	Equivalence of Frequency Data (AE)	37
	Equivalence of Training Adequacy Data (AE)	38
	Equivalence of Proficiency Data (AE)	38
	JKT versus SI (AE)	39
	AE Subarea Correlations	39
	JKT versus Q Method	40
	Method Comparability (BT)	40
	AE versus BT Data	41
	Overall Scale Means	41
IV	RESOURCE REQUIREMENTS AND DISCUSSION	43
	Resource Data	43
	AE	43
	BT	45
	Cost Per Instrument	46
	PEP Costs	46
	Estimated Appraisal Costs	47
	Discussion	49

TABLE OF CONTENTS (continued)

Section		Page
V	DISCUSSION OF ADMINISTRATIVE/TECHNICAL PROBLEMS	53
	Fleet Coordination	53
	Study Samples	53
	Facilities	54
	Privacy Act	54
	Questionnaire Return Rates	54
	BT Data Inconsistencies	56
	NOTAP Participation	57
	Need for Single Point Control	58
VI	CONCLUSIONS AND RECOMMENDATIONS	59
	Conclusions	59
	Method Equivalence	59
	Source Equivalence	59
	Independence of Scales	59
	Resource Comparisons	60
	Recommendations	60
REFERENCES		63
APPENDIX A	Specific Job Tasks Rated by AEs	65
APPENDIX B	Specific Job Tasks Rated by BTs	73
APPENDIX C-1	Tasks Rated Within Specified Subcategories of Frequency Scale by AEs	81
APPENDIX C-2	Tasks Rated Within Specified Subcategories of Training Adequacy Scale by AEs	87
APPENDIX C-3	Tasks Rated Within Specified Subcategories of Proficiency Scale by AE Supervisors	93
APPENDIX D-1	Tasks Rated Within Specified Subcategories of Frequency Scale by BTs	97
APPENDIX D-2	Tasks Rated Within Specified Subcategories of Training Adequacy Scale by BTs	103
APPENDIX D-3	Tasks Rated Within Specified Subcategories of Proficiency Scale by BT Supervisors	109
APPENDIX E	Scale Independence: Comparative Data	113
APPENDIX F	Overall Scale Means and Standard Deviations	115

LIST OF ILLUSTRATIONS

Figure		Page
1	Sample Page of BT Q/SI Data Collection Instrument.	17

LIST OF TABLES

Table		Page
1	Sampling Plan for Administration of Mailout Questionnaires (Q), Structured Interviews (SI), and Job Knowledge Tests (JKT) to AE "A" School Graduates and Supervisors.	14
2	Functional Subareas Covered by Instruments.	18
3	Planned Comparisons for Determining Equivalence of Methods and Data Sources.	27
4	Intercorrelations of Mean Scale Ratings Given by AEs on the Q and SI Instruments by Graduates (G) and Supervisors (S) . . .	32
5	Correlations Between AE Graduates JKT Scores and Ratings of Training Adequacy and Proficiency on the Structured Interview .	33
6	Correlations Between AE JKT Scores and SI Mean Scale Values on Corresponding Items.	33
7	Correlation Values Obtained Between JKT Subarea Scores and SI Subarea Mean Scale Ratings (AE).	34
8	Intercorrelations of Mean Scale Ratings Given by BTs on the Q and SI Instruments by Graduates (G) and Supervisors (S) . . .	35
9	Correlations Between BT Graduates JKT Scores and Ratings of Training Adequacy and Proficiency on the Structured Interview .	36
10	Correlations Between BT JKT Scores and SI Mean Scale Values on Corresponding Items	36
11	Correlation Values Obtained Between JKT Subarea Scores and SI Subarea Mean Scale Ratings (BT)	37
12	Aviation Electrician's Mate Cost Data	43
13	Boiler Technician's Cost Data	45
14	Actual and Projected Cost Per Instrument.	46
15	Comparative Costs of Data Collection.	48
16	Questionnaire Return Rates for AE and BT Graduates and Supervisors	55
C-1	Number of Tasks Rated Within Specified Frequency Subcategories by Graduates (G) and Supervisors (S) (AE)	81

LIST OF TABLES (continued)

Table		Page
C-2	Number of Tasks Rated Within Specified Training Adequacy Subcategories by Graduates (G) and Supervisors (S) (AE)	87
C-3	Number of Tasks Rated Within Specified Proficiency Subcategories (AE).	93
D-1	Number of Tasks Rated Within Specified Frequency Subcategories by Graduates (G) and Supervisors (S) (BT)	97
D-2	Number of Tasks Rated Within Specified Training Adequacy Subcategories by Graduates (G) and Supervisors (S) (BT)	103
D-3	Number of Tasks Rated Within Specified Proficiency Subcategories (BT).	109
F-1	Overall Scale Means and Standard Deviations	116

SUMMARY

The Chief of Naval Technical Training (CNTECHTRA) initiated a Pilot Evaluation Project (PEP) to evaluate procedural concepts and to define components of a training appraisal subsystem for subsequent use in the systematic collection of training feedback data from the fleet. The Training Analysis and Evaluation Group (TAEG) was tasked by the Chief of Naval Education and Training (CNET) to assist CNTECHTRA in the PEP study.

A principal objective of the PEP was to evaluate the relative merits and optimum circumstances for the use of three different methods of obtaining training feedback data from the fleet. The methods evaluated were a mailout questionnaire (Q), a structured interview (SI) procedure, and a job knowledge test (JKT).

Training Analysis and Evaluation Group tasking required the development of Q and SI instruments for collecting feedback data from recent school graduates and their respective supervisors. Both instruments were designed to solicit information concerning:

- the frequency with which a given task(s) is performed on the job by recent graduates,
- the adequacy of school training for the task(s), and
- ratings by supervisors of the proficiency with which recent graduates perform job task(s).

The Central Test Site (CTS) for the Personnel and Training Evaluation Program (PTEP) was tasked to develop and administer JKTs. These were designed to assess graduates' factual knowledge of operations required to perform various specific job tasks. Members of PTEP also conducted SIs at various fleet units. The TAEG administered the mailout Qs.

The TAEG was also tasked to analyze data relevant to an evaluation of the three methods. This report presents the results of the requested comparative evaluations. The methods were evaluated on the basis of the data generated by each, resource requirements associated with the use of each, and administrative and technical problems peculiar to a method's use.

To compile information for the method comparisons, training feedback data were collected via the three methods in the Atlantic Fleet from Aviation Electrician's Mates (AE) and Boiler Technician (BT) 1200 PSI "A" school graduates and fleet supervisors of these graduates.

The results of the study support conclusions that:

- Equivalent frequency of task performance data can be obtained by either the Q or the SI method.
- The Q and SI yield equivalent ratings of training adequacy.

- Assessment of proficiency can be accomplished using either the Q or SI method.

- The Q/SI and JKT methods do not produce equivalent information about training effects.

In brief, the Q and the SI methods used in this study are fully interchangeable for data collection. The JKT method produces different information.

Analyses performed to assess the equivalence of data sources support the general conclusion that comparable frequency of task performance and training adequacy information can be obtained from either graduates or supervisors.

The results of analyses of resource data demonstrate that, overall, the Q method is least expensive for collecting data. The SI method is more costly than the Q method because of the one-on-one assessment situation. Primarily, because of high developmental costs, the JKT is the most expensive technique. These costs are directly proportional to the number of tests administered with costs of a given JKT decreasing as the number of tests increase. For the same numbers of individuals "tested," total SI costs would increase because of the two-person situation. Mailout Q costs will vary as a function of postage and reproduction costs. The final choice of a data collection method, however, should not be based solely on resource considerations. As discussed in the text, the instrument used must consider fully the information required from an appraisal effort.

Based on the results of this study the following recommendations are made for method choice:

1. Low cost questionnaires of the type used in this study are recommended for use as screening devices to determine the presence (or absence) of deficiencies in the quality of school training.

2. Structured interview procedures are recommended when more detailed information is required concerning the specific nature of training deficiencies.

3. Job knowledge testing should be restricted to the purposes for which it is intended; i.e., to assess trainee possession of factual knowledge of job/equipment operations and repair. In its current state of development and application, it seems inappropriate for field use as a technique for identifying deficiencies in school training.

A number of detailed recommendations for conducting training appraisals in the future are provided in section VI of this report.

TAEG Report No. 64

SECTION I

INTRODUCTION

BACKGROUND

The Training Analysis and Evaluation Group (TAEG) was tasked by the Chief of Naval Education and Training (CNET)[1] to support the Chief of Naval Technical Training (CNTECHTRA) in the conduct of a Pilot Evaluation Project (PEP). The PEP was concerned with the evaluation of procedural concepts and the assessment of various methods of obtaining training feedback data for the Naval Education and Training Command (NAVEDTRACOM). The major objectives of the PEP were to:

1. field test a Fleet Feedback Data Collection Group (FFDCG) concept with focus on the resource requirements for implementation,

2. develop procedures for the evaluation of a total training pipeline,

3. evaluate the relative merits and optimum circumstances for use of three different methods of gathering feedback, and

4. use the findings to recommend a standardized training appraisal subsystem for CNET.

Various Navy organizations participated in the PEP study. The TAEG was tasked to develop questionnaire (Q) and structured interview (SI) instruments and procedures for their administration. The TAEG was also assigned responsibility for data analysis. At the time of this tasking the basic conceptual design of the study had been established by CNTECHTRA. Similarly, the management structure for the conduct of the study was also defined. The Central Test Site (CTS) for the Personnel and Training Evaluation Program (PTEP) was designated as an FFDCG with responsibility to effect fleet interface/coordination to conduct JKTs and SIs at fleet locations. The CTS for PTEP was also tasked to develop job knowledge tests (JKT) for collecting feedback data. The Navy Occupational Task Analysis Program (NOTAP) Department of the Navy Occupational Development Analysis Center (NODAC) was tasked to provide job task data printouts and computer support for the processing of survey data. Overall management responsibility for the PEP was delegated to CNTECHTRA (Code 016). Chief of Naval Education and Training (now Code 015) provided policy guidance and command liaison functions.

Experiences with the fleet in gathering feedback, identification and comparisons of differential resource requirements, technical and administrative problems encountered, and analyses of the data gathered by the three different data collection methods were needed to provide information relevant to the PEP objectives. To acquire this needed experience and information base, it was necessary to select target "A" schools/courses about which fleet feedback information could be obtained. The Aviation Electrician's Mate (AE) "A" School and the Boiler Technician (BT) "A" 1200 PSI Course were selected.

[1] CNET ltr Code N-5 of 12 Nov 1975

Corollary objectives of the PEP study were to obtain and report feedback information to these schools which could be used to assess needs for curriculum revisions. Summaries of the feedback data collected during the PEP were prepared and provided the schools concerned via the following reports:

1. Central Test Site for the Personnel and Training Evaluation Program, PTEP Evaluation 8-77, Aviation Electrician's Job Knowledge Test Report. October 1977. Central Test Site for the Personnel and Training Evaluation Program. Naval Guided Missile School, Dam Neck, VA 23461

2. Hall, E. R. and Denton, Carol F. Aviation Electrician's Mate "A" School Training Assessment Data, TAEG Technical Memorandum 77-9. November 1977. Training Analysis and Evaluation Group, Orlando, FL 32813

3. Denton, Carol F. and Hall, E. R. Boiler Technician "A" School (1200 PSI) Training Assessment Data, TAEG Technical Memorandum 78-1. July 1978. Training Analysis and Evaluation Group, Orlando, FL 32813

4. Central Test Site for the Personnel and Training Evaluation Program, PTEP Evaluation 7-78, Boiler Technician (1200 PSI) Job Knowledge Test Report. December 1978. Central Test Site for the Personnel and Training Evaluation Program. Naval Guided Missile School. Dam Neck, VA 23461

PURPOSE

The purpose of this report is to present an analysis of data obtained via the three data collection techniques. Accordingly, this report primarily addresses the evaluation of the relative merits and optimum circumstances for use of the three different methods of gathering feedback. Where possible, and justifiable, information bearing on other PEP objectives is reported. However, no detailed discussions of, or conclusions with respect to, those objectives are presented.

ORGANIZATION OF THE REPORT

The remainder of this report is presented in five sections and six appendices. Section II provides details of the methodological and experimental design of the study as well as descriptions and details of resource accounting and tracking of administrative and technical problems. Section III presents and discusses results of analyses which have significance for the comparative evaluation of methods. This technical section presents the statistical details upon which study conclusions are based. Section IV contains an analysis of resources consumed in the collection of feedback information via the three methods and discusses implications for future feedback efforts. Administrative and technical problems encountered during the PEP are addressed in section V. The final section of the report presents conclusions and recommendations. Appendices A and B list job tasks evaluated by AEs and BTs, respectively. Appendices C and D summarize Q and SI evaluation results for AEs and BTs. Appendices E and F present data pertaining to statistical comparisons.

SECTION II

TECHNICAL APPROACH

This section presents details of the technical approach. A discussion of the rationale underlying the study is presented first. This is followed by descriptions of the methodological and analytical design of the study. Procedures used for identifying/tracking resource requirements and administrative and technical problems incidental to aspects of the PEP are also described. The discussions presented below are deliberately brief. Only information considered necessary for understanding, interpreting, and properly evaluating the results of comparing the three different data collection methods is presented.

STUDY RATIONALE AND LIMITATIONS

Meeting objective 3 of the PEP required the collection of information about the three data gathering methods. Specific information was required concerning the feedback data obtainable by each method, the resources required to obtain those data, and any technical or administrative problems peculiar to their use. In short, information was needed concerning the advantages or disadvantages of using each method and for identifying any limitations on their use. It was intended that the study produce information for future decisionmaking concerning the selection of a best method of obtaining feedback from the fleet for some given set of circumstances.

From the TAEG viewpoint, the primary and proper focus of a comparative study of this nature should be on the data obtainable via the different methods. In this regard, two issues are of concern:

1. Which method(s) generates relevant and valid information?

2. Do the methods produce equivalent information?

Answers to the first question are essential to determining the value of data for identifying training deficiencies in terms conducive to their correction. Unfortunately, no provision was made in the original study design for assessing method validity. Consequently, the TAEG team addressed the second question regarding equivalency of methods. Findings that the methods were equivalent in terms of data produced would support decisions concerning their interchangeability for future data collection efforts. If equivalent, methods could be selected on the basis of cost and/or ease of use factors. For these determinations, information was also sought concerning resources required for use of the methods and note was made of technical and administrative problems peculiar to the use of each method.

METHODOLOGICAL DESIGN

The following paragraphs describe the plans and methods employed to obtain, summarize, and analyze feedback data.

PEP STUDY PLAN. As mentioned earlier, the basic study design was conceived by CNTECHTRA (016). Operating within the conceptual framework imposed, TAEG added design features where possible to collect information bearing directly on specific questions concerning method equivalency (e.g., matching supervisors with graduates, adding a proficiency scale to the Q/SI instruments to provide a basis for evaluating the JKT).

The focal point of evaluation was defined as an individual graduate of an "A" school. The aggregation of information about/from individual graduates provided the basic data pool for addressing the issue of equivalency of data. (Records of resource requirements and problems of using each method would provide other necessary information.) Plans were to obtain feedback concerning both Atlantic Fleet and Pacific Fleet graduate-assignees. However, difficulties and delays in securing Atlantic Fleet approval and support significantly extended the study schedule. To avoid the further delays that might occur by attempting to obtain Pacific Fleet approval, the decision was made to confine the survey to the Atlantic Fleet.

For the study, names of recent "A" school graduates were obtained from the AE and BT schools, respectively. Samples were drawn from these identified groups. Each sample was further subdivided to collect feedback via the three different methods as explained below.

AE Plan. For the AE portion of the study, names and current duty station addresses of all individuals who had graduated from the "A" school during the period 11 November 1975 to 2 June 1976 were obtained from the school. The original PEP study design called for assessing two groups of AE "A" school graduates. One group was composed of 200 graduates currently working in the fleet, the other of 200 graduates undergoing advanced training within the NAVEDTRACOM. The AE study plan is shown in table 1.

TABLE 1. SAMPLING PLAN FOR ADMINISTRATION OF MAILOUT QUESTIONNAIRES (Q), STRUCTURED INTERVIEWS (SI), AND JOB KNOWLEDGE TESTS (JKT) TO AE "A" SCHOOL GRADUATES AND SUPERVISORS

SOURCE OF DATA	Q		SI		JKT	
	Grad	Supv	Grad	Supv	Grad	Supv
Fleet	125	125	50	50	75	0
NAVEDTRACOM	125	25	50	15	75	0

Questionnaires were to be mailed to 125 graduates in the fleet with the same number mailed to their supervisors. Since any one supervisor could have several of the graduates, this would require that he complete as many forms as he had graduates (i.e., one for each graduate). Fifty graduates in the fleet were to be given SIs with appropriate supervisors also providing 50 SIs--one for each graduate supervised. It was also specified that the 50 graduates

given SIs would be given the JKT so that test scores could be directly compared to the SI results. Job knowledge tests were to be administered to an additional 25 AE graduates. No individuals in the Q group were to be given any other treatment. Under this design, the maximum number of instruments that would be available for analysis from the fleet group would be: 250 Qs (125 graduates plus 125 supervisors), 100 SIs, and 75 JKTs.

Table 1 also shows the plan for data collection within the NAVEDTRACOM. Individuals in Naval Air Maintenance Training Detachments (NAMTRADET) were the intended subjects. It was hoped that results equivalent to those from the fleet could be obtained from a school population. This finding would obviate the necessity for going to the fleet for feedback data for future evaluation studies. This design feature would also provide information bearing on objective 2 of the PEP dealing with procedures for pipeline evaluation. However, it became necessary to abandon this plan when it was learned that NAMTRADET training is irregularly scheduled and small numbers of students typically are involved in classes. Hence, an unduly long time period would be required to collect necessary data. Also, since students have virtually no job experience, it would not be appropriate to request their comments on the adequacy of "A" school training for the job. Being job-oriented, the survey items (i.e., tasks) were also deemed not appropriate for assessing graduates' preparation for additional/advanced training (i.e., for progression from one learning environment to another).

BT Plan. The BT study plan was highly similar to the AE plan. Individuals graduating from the 1200 PSI course during the period of June 1976 to February 1977 formed the basic group of interest.

The BT plan called for administering Qs to 125 recent graduates of the BT "A" course (1200 PSI) and their respective supervisors. An additional 50 graduates and their supervisors were scheduled for SIs. Again, the graduates given SIs were tested by the CTS for PTEP using JKTs developed specifically for the PEP study. One modification to the design was made, however, to permit a statistical comparison that was not provided for in the AE data collection program. The design change made involved adding a group of 40 graduates who would be given a JKT plus being asked to complete Qs. Their corresponding supervisors were also given Qs to complete. This change was made to provide a basis for comparing Q and SI responses of groups of graduates on the basis of a common JKT experience.

In summary, the study design was considered adequate to answer questions concerning the equivalency of the three methods. Unfortunately, however, the issue of validity was inadequately considered in CNTECHTRA's original design. This omission limits the conclusions that can be drawn with respect to the question of "best" method.

DESCRIPTION OF THE INSTRUMENTS. The three data collection instruments are described briefly below. Only details necessary to understand the results of the methodological comparisons are presented. Greater detail may be found in the four published reports previously cited on page 12.

Q and SI Data Collection Instruments. The Q and SI instruments used for data collection were identical in format. Instruments used with graduates, however, contained only two scales while three scales were used on the supervisor forms. Figure 1 shows a sample page of a form used with BT supervisors. The left hand column listed specific job tasks for either an AE or a BT. (A complete list of the job tasks evaluated is given in appendix A for the AE rating and in appendix B for the BT rating.) Subsequent columns provide space for respondents to rate:

- the frequency with which various tasks are performed on the job by a particular school graduate,
- the adequacy of school training for the performance of a particular (required) task, and
- the proficiency (supervisors only) with which a given school graduate performs the given task.

Ratings on these scales are made by selecting an appropriate/desired option (1, 2, 3, 4, or 5) from the boxes below the column headings. Option choice is indicated by circling the appropriate number in the proper column on the row listing a given task.

JKT. The JKTs used for AE and BT data collection were developed specifically for the PEP by the CTS for PTEP. The JKTs were designed to assess the job knowledge of recent (1 year or less) graduates of the two "A" schools. Each JKT was divided into two parts: knowledge and skill. The knowledge parts of each test consisted of 100, four alternative, multiple choice test items. These were designed to test knowledge of facts specific to each rating. For diagnostic purposes the knowledge test items were subdivided into eight functional areas. This was done to determine if graduates' strengths or weaknesses fell into defined curricular areas. These are shown in table 2. The skill parts of each test consisted of two paper and pencil, simulated, troubleshooting exercises. These exercises presented an equipment failure which the examinee was required to analyze. However, only the knowledge test items were used for comparing methods of data collection.

Reliability of the JKTs was determined by the Kuder-Richardson Formula 20 (KR-20). An explanation of this technique is provided in Guilford (1954). This statistic provides a measure of the homogeneity of a test; i.e., the extent to which each item measures what is measured by all other items. For the AE JKT, reliability was determined over a base of 85 graduates. The correlation coefficient for total test reliability was determined by the KR-20 technique to be equal to .75 with a standard error of 4.29 (see Guilford, 1954). The KR-20 reliability coefficients for the subparts of the AE JKT are shown in table 2.

BT test reliability was computed over a base of 51 graduates. These were the same graduates who had also been given an SI. The correlation coefficient for reliability of the total BT test as determined by KR-20 was equal to .68

BEST COPY AVAILABLE

PAGE 4	FREQUENCY OF TASK	ADEQUACY OF BT "A" SCHOOL TRAINING FOR THIS TASK	PROFICIENCY AT DOING TASK UPON ARRIVAL FROM "A" SCHOOL
Listed below are tasks which presently receive at least some emphasis in BT "A" School. Please rate each task on the scales at the right by circling the most appropriate number. Please feel free to also include your reasons for your rating and/or specify recommendations for training on this task. Your comments may be written in any available space on the front or back of this page or on the "Comments" sheets at the back of this booklet.	1. Never performed.* 2. Seldom performed or only in emergencies. 3. Performed monthly. 4. Performed weekly. 5. Performed daily. * ADEQUACY AND PROFICIENCY SCALES may be skipped if task is never performed.	1. Task requires much more emphasis in "A" School. 2. Training less than adequate for rank, increase emphasis in "A" School. 3. "A" School training adequate for rank. 4. Training more than adequate for rank, reduce emphasis in "A" School. 5. Greatly reduce or eliminate "A" School training for this task.	1. Extremely limited in part owing rank, cannot do without direct supervision. 2. Performs task with difficulty, requires general supervision. 3. Partially proficient, requires intermittent supervision. 4. Performance is adequate without supervision. 5. Performs task with ease, highly proficient.
1. (G15) USE ENGINEERING OPERATIONS SEQUENCING SYSTEM MANUAL (EOSS)	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
2. (OC7) PERFORM BOILER PRELIGHT-OFF INSPECTIONS	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
3. (P29) LINE UP/SECURE MAIN FEED RECIRCULATION SYSTEM	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
4. (S1) LINE UP FUEL OIL HEATER DRAINS IN ACCORDANCE WITH EOSS	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
5. (P36) LINE UP/SHIFT TO-FROM LP/HP DRAIN SYSTEMS	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
6. (P57) LINE UP MACHINERY COOLING WATER REDUCING STATION	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
7. (S2) LINE UP FRESH WATER DRAIN COLLECTING TANK SYSTEM IN ACCORDANCE WITH EOSS	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
8. (S3) LINE UP FRESH WATER DRAIN COLLECTION SYSTEM TO DEAERATING FEED TANK IN ACCORDANCE WITH EOSS	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5

Figure 1. Sample Page of BT Q/SI Data Collection Instrument

TABLE 2. FUNCTIONAL SUBAREAS COVERED BY INSTRUMENTS

Subarea	Rel.	AE Functional Areas	Subarea	Rel.	BT Functional Areas
C	.04	Wiring	C	.49	Boiler Operation
D	.36	Basic Troubleshooting	D	.20	Boiler Maintenance
E	.50	Supply and Administration	E	.46	Boiler Cleaning
F	.43	Aircraft Instruments	F	.59	Casualty Control
G	.36	Engine Performance/ Warning Systems	G	.23	Administration/General Military Training
H	.45	Electrical/Electronic Systems	H	.13	Valves
I	.35	Auto Pilot Systems	I	.04	Fire Room Maintenance
J	.22	Scheduled Maintenance	J	.25	Auxiliary Maintenance

(with a standard error of 4.92). The KR-20 BT subpart reliabilities are also shown in table 2. Item statistics for both tests may be obtained from the CTS for PTEP.

CORRESPONDENCE OF INSTRUMENT ITEMS. As previously noted, the statements on the Q/SI instruments identified tasks performed on the job by AEs and BTs, respectively. JKT items were written to assess graduates' factual knowledge of job information required to perform such tasks. All 100 JKT test questions corresponded to tasks in the Q/SI listing. For the BTs, 95 of the 100 test items corresponded to tasks in the Q/SI listing. In a number of instances, several test items were written for a particular task statement. This reduced the number of pairs of matched items to 47 for the AE rating and 21 for the BT rating. This overlap between instruments was designed to permit direct comparisons of measured levels of graduates' knowledge with perceptions of the adequacy of "A" school training and supervisors' judgments of proficiency.

INSTRUMENT DEVELOPMENT. The procedures used in developing the feedback instruments are described below.

Q/SI Instruments. The basic format of the Q/SI instruments was taken from Dyer, Ryan, and Mew (1975). The frequency of task performance and adequacy of training scales were taken from the procedure recommended in that report. The proficiency scale was developed for the present study. It was added by TAEG to the Q and SI instruments to provide a criterion for estimating the validity of each JKT.

Task statements contained in the instruments were developed from analyses of the NOTAP job task inventory for each rating. From these, school personnel identified job tasks which received some level of training in the school curriculum. All 124 tasks listed on the AE forms had received some training emphasis. A number of tasks for which the BT school did not provide training were included on that survey format, however. These were tasks which the NOTAP task inventory identified as being performed by substantial numbers (i.e., 20 percent or greater) of BT3s in the fleet. It was reasoned that data collected on these tasks would support decisions concerning whether to include them in future training. These "nonschool-trained" BT tasks are identified in appendix B. By coincidence the BT Q/SI instruments also contained 124 task statements.

Prior to use in the field, the AE Q and SI forms were critically reviewed by key staff and faculty members of the AE "A" school and by the staff of CNTECHTRA (Code 016) and CNET (Code 015). Appropriate revisions were made from their comments. In addition to these reviews, the instruments were field tested on 21 AE-rated personnel. This group (composed of five Chief Petty Officers, four First Class, five Third Class, and seven Airmen) represented several aircraft maintenance shops, various aircraft, and different levels of skill. The results of the field pretest indicated that the instruments were suitable for the subject population and that they could provide useful training effectiveness data.

Prior to their use in the field, the BT Q and SI forms were critically reviewed by key staff at the Propulsion Engineering School (Great Lakes) and by faculty members of the Main Propulsion Maintenance (BT) School (Norfolk). CNTECHTRA (Code 016) personnel, the BT School Training Program Coordinator (TPC), and CNET (Code 015) staff also reviewed the materials. Appropriate revisions were made from comments received. In addition to these reviews, the instruments were field tested on 13 BT-rated personnel. Seven of these were BT3s undergoing advanced training at the Air Conditioning and Refrigeration (AC&R) School (Norfolk). Five (E-5/E-6) were instructors at the AC&R School, and one E-8 was assigned duty at the Recruit Training Command, Orlando. The results of the field pretest indicated that the instruments were suitable for use with the subject population and that they could provide useful training effectiveness data.

JKT Development. Significant features of development of the JKTs are outlined below. This information was summarized by TAEG from inputs provided by the CTS for PTEP.

AE JKT. An AE JKT development group composed of PTEP personnel and senior enlisted personnel (E-6 and up) from the NAMTRADET, Oceana, analyzed the NOTAP AE job task analysis data. The purpose of this analysis was to identify AE performances required/common across the AE rating; i.e., task performance not specific to any given aircraft. This group wrote a series of preliminary test items covering the job knowledges required for the performance of the selected tasks. Other test items were given PTEP for inclusion in the JKT by the AE "A" school. A preliminary version of the test was developed by PTEP from these two sources.

The preliminary item pool was then pretested to obtain information for selecting items for retention in a final version of the test. The preliminary version of the test was administered to 36 AEs in pay grades E-2 through E-7. Standard PTEP data procedures and item analysis techniques were used to compute item statistics/characteristics. Content validity of the JKT was established by the PTEP.

For the items, validity indices (VI), difficulty indices (DI), and reliability values were computed. These item statistics identified 10 test questions with unsatisfactory values. These 10 items were reviewed by subject matter experts (SME) to determine that each item reflected skills and knowledges employed by AEs in the performance of their duties. For this, SMEs rated the items 1, 2, 3, or 4 as to their relevancy. As a result of this review, these 10 items were deemed nonrelevant and were replaced with other items.

AE "A" school personnel reviewed and endorsed the test prior to its use in the field. School personnel also determined that all test items were based on tasks for which the school had given training. The final test was nonaircraft-specific.

BT JKT. Except as noted below, the BT JKT was developed in the same manner as the AE JKT. Tasks were selected for test item development from the NOTAP BT Job Task Inventory (JTI). Personnel from PTEP used the principal criterion

for task selection as those tasks performed by 20 percent (or more) of BT3s. Secondary criteria were the degree of member involvement and time spent performing the task. The components of these secondary criteria were considered in selecting tasks. Degree of member involvement, for example, contains the subcategories of "perform," "assist," "supervise." Tasks were selected from this category with the heaviest weighting given to "perform," second to "assist," and third to "supervise." Selection from the "time spent performing the task" category gave the heaviest weightings to tasks requiring the greatest time.

This selection process yielded an initial task pool of 100 items which was subsequently reduced by inspection. This yielded 64 tasks considered to be of possible testing interest. The 64 tasks were then compared to Personnel Advancement Requirements (PAR) and to Personnel Qualification Standards (PQS). The amenability of tasks to evaluation by knowledge questions was also considered in this second selection process. This analysis reduced the task pool to 34. These 34 job tasks then provided the basis for the BT JKT.

Initially, 189 four alternative multiple choice test items were developed by the CTS to assess the knowledge which supported the performance of these 34 job tasks. The items were critically reviewed for PTEP by personnel from the Main Propulsion Maintenance (BT) School at Norfolk for technical accuracy/correctness, NOTAP task identifier support, universality (i.e., that the items sampled knowledges required across the BT rating and were not specific to any particular ship), and content validity. Thirty-six items were totally rejected during this review. From the remaining items, these SMEs selected the 100 "best" based on criticality of the task involved and the applicability to the Fireroom Upper Level, Lower Level, and Messenger watch stations.

The preliminary version of the test was then administered to a pilot group of 32 BTs. This consisted of:

- seven school staff personnel (two E-8's, four E-7's, and one E-6)
- sixteen fleet experienced BTs (one E-6, seven E-5's, and 8 E-4's)
- nine non-BT "A" school graduates (one E-3 and eight E-4's).

Item analyses were performed on the pretest results. The item statistics revealed 36 items that did not meet preset (arbitrarily chosen) item difficulty values (DI) (0.4) and/or item discrimination (VI) (0.3) criteria. These were reviewed by SMEs who considered 27 of the items still to be relevant and valid. These items were retained in the final version of the test. The remaining nine were either replaced or rewritten. The final version of the BT test was nonplant and nonship specific. It also contained a number of items reflecting tasks that were not taught in the BT curriculum.

ADMINISTRATION OF INSTRUMENTS. Procedures used to administer the three data collection instruments are described below.

Questionnaires. The AE questionnaires were mailed on 27 January 1977 to the commanding officers of squadrons to which the AE graduates of interest were assigned. BT questionnaires were mailed in two lots--one on 2 September 1977 and the other on 5 December 1977. Questionnaire packages were transmitted via a Commander in Chief, U.S. Atlantic Fleet (CINCLANTFLT) letter[2] requesting the personal support and cooperation of commanding officers in ensuring Q completion and return. Commanding officers were requested to distribute the packages to affected individuals (i.e., graduates and their supervisors) under their command.

AE respondents were instructed to return completed forms directly to the TAEG in preaddressed envelopes provided. Since names and social security numbers were requested, it was originally interpreted that the Qs were subject to the provisions of the Privacy Act; thus, completion was voluntary. Subsequently, it was determined that this act did not apply to military training surveys.[3] Respondents were instructed to return the completed forms within 2 weeks. Because of an initially low return rate, an additional 25 sets of AE questionnaires were mailed in March 1977. These were mailed to insure that a sufficiently large number of returns would be available for subsequent statistical analyses.

In an attempt to improve return rates, an administrative change was made for return of the BT Qs. Again, Q packages were transmitted under the CINCLANTFLT cover letter. However, this time respondents were instructed to return the completed Qs to their commanding officers for transmittal to the TAEG.

Initially, BT Q return rates from the ships were low. As a followup procedure, the Commander, Naval Surface Force, U.S. Atlantic Fleet (COMNAVSURFLANT) sent a message[4] to all affected ships requesting a one-time status report covering their degree of participation in the effort. Also, in an attempt to obtain a larger return rate, commanding officers receiving Qs in the second mailing were requested to have any BT "A" Course (1200 PSI) graduate (and his supervisor) under the command complete the Qs if the graduate of interest was no longer available.

Thus, in the BT case, completion and return of the Qs had the appearance, at least, of being mandatory rather than voluntary as in the AE situation. This administrative change may have had certain undesirable side effects. This is discussed in section V.

SI/JKT General Procedures. Personnel from PTEP conducted all SIs for both ratings and also administered all JKTs. Selected units were contacted via

[2] CINCLANTFLT ltr ser 6950 of 15 Nov 1976

[3] SECNAVINST 5211.5A, OP09B1, 8 Oct 1976

[4] COMNAVSURFLANT msg 150223Z Nov 1977

telephone to explain the purposes of the PEP study and the assistance desired. These calls were placed to commanding officers or executive officers. In the BT case, the Chief Engineer and/or the Main Propulsion Assistant (MPA) may also have been contacted if and when appropriate.

After this initial contact, the PTEP team went to the appropriate ship (squadron in the AE case), presented a copy of the CINCLANTFLT letter endorsing the project, and briefed the individuals responsible on the purpose of the project and cooperation desired. Reports concerning the JKT results that would be given back to the ship/squadron were also described. The PTEP team then administered JKTs and/or conducted SIs. In the Charleston (South Carolina) area, PTEP spaces were used for testing and interviewing. In other locations, available facilities were used.

SIs. AE SIs were conducted under PTEP supervision in the Norfolk (Virginia) and Jacksonville (Florida) areas between 11 April 1977 and 15 May 1977. BT interviews were conducted in the Norfolk and Charleston areas between September 1977 and February 1978.

Eight senior enlisted Naval personnel served as interviewers for the AE portion of the effort. Seven were AEs assigned temporary duty (from NAMTRADETs and operational squadrons) to assist in the effort. The eighth, an ETCM, was assigned permanent duty at the CTS for PTEP. BT interviews were conducted by seven senior enlisted Naval personnel. None were BT rated. All were assigned permanent duty at PTEP. Some were attached to the CTS at Norfolk (Dam Neck), others to the PTEP Charleston Detachment.

The PTEP interviewers were given procedural guidance by the TAEG for conducting the SIs. Forms were also provided on which resources expended to gather the data could be recorded. Structured interviews were conducted in a face-to-face, one-on-one setting. During the SI, the interviewer read the task statements to the respondent. The subject responded verbally and the interviewer recorded his responses (i.e., scale choices) on the SI form. Respondents were provided opportunity to comment on questions or to explain answers. Graduates were interviewed first. Supervisors of the graduates were interviewed within 1 to 14 days after the graduate was interviewed. For the AEs, SIs were held before the graduates were given JKTs. This was changed for the BT portion of the study so that approximately half of the graduates were given the JKT before being interviewed; the other half were interviewed first and then given the JKT.

In both the AE and BT cases plans to use preidentified graduates exclusively could not be carried out. Due to deployment schedules, some graduates could not be contacted within the time available for the study; others were on leave or had been transferred to new commands. Still others simply never arrived at the Command identified by the schools. Thus, "available" BT and AE graduates and their supervisors were interviewed as well as individuals from the planned groups.

JKTs. JKTs were administered to groups of AEs and BTs. Typically, 10 to 15 individuals were in each group; but 35 BTs were tested in one instance. AEs

were allotted 2 hours and 20 minutes to complete both portions of the JKT (i.e., knowledge and skill); BTs were allotted 2 hours and 10 minutes for both parts of the test. On the average, 40 minutes more were required for each testing session to brief examinees, explain the Privacy Act, and otherwise prepare for the test.

DATA HANDLING/PROCESSING. The mechanics of data handling are described below.

Q/SI Forms. Questionnaires were returned to the TAEG in the manner previously specified. Structured interviews were collected by the PTEP interviewers and mailed in groups.

For AEs, the data were recorded directly onto the Q and SI booklets by respondents and interviewers and were transferred to NOTAP machine processing booklets by TAEG personnel. These were mailed to NODAC for processing by the NOTAP Department. Raw data (i.e., individual task ratings) were also entered into a WANG programmable calculator (System 2200 VP) at TAEG for later processing and preparation of other data summaries.

For a number of reasons, NOTAP facilities were not used for processing BT Q and SI data. These included:

- the extra man-hours required to transfer the data from the Q and SI booklets onto the NOTAP forms,
- the inconvenience and difficulties of coordinating report preparation via the telephone and through the mail, and
- long turn around times resulting from the above and also from the fact that NOTAP data are processed on a computer which is available only on weekends.

For these and other reasons discussed in section V, more economical and efficient data handling procedures were followed for the BT information. In this case, data recorded on the Q and SI booklets were entered directly into the WANG calculator. These data were then transferred to a tape for processing on more powerful computational equipment. Summaries of the Q and SI data were prepared for school use in assessing the BT 1200 PSI curriculum. Some were processed within the TAEG. Others were prepared by the Navy Regional Data Automation Center (NARDAC) Detachment, Orlando, according to TAEG specifications. Data summaries were provided the schools as stated in the reports previously cited (see page 12).

The summaries prepared for school use in assessing the current curriculum consisted largely of the mean ratings given the job tasks on each of the three scales: frequency, training adequacy, and proficiency. Mean scale ratings were computed for various combinations of the two classes of respondents (i.e., graduates and supervisors) and for the two instruments (i.e., mailout questionnaires and structured interviews). All means were computed conventionally as the sum of the scale values given by individuals within a group

divided by the number of individuals within that group. Various combinations of data (e.g., by class of respondent, by type of instrument) were made.

Appendices C and D present selected summaries of AE and BT data, respectively. These summaries show the numbers of tasks rated within specified subcategories of the frequency, training adequacy, and proficiency scales. The specific tasks falling within these subcategories are also identified. These summaries are presented both for information regarding how the job tasks were rated on the Q and SI instruments and also for facilitating the interpretation of the results of the analyses done to compare the methods.

JKTs. Graduates who took JKTs recorded their answers directly onto machine readable standard answer sheets. These answer sheets were optically scanned at the CTS for PTEP and the results were fed directly (via telephone line) to an IBM computer at the Polaris Missile Overhaul Facility at Charleston. Data were processed in accordance with standard PTEP data processing programs resident in the computer. Item statistics were computed. On these bases some items were eliminated (as invalid) from further statistical analysis. Standard PTEP data summaries were prepared at Charleston as was a Test Analysis Manual (TAM). These were transmitted back to the CTS via teletype. Reports of examination results of individuals assigned to particular commands were also prepared and mailed back to the participating squadrons/ships. Data interpretation and report writing were accomplished at the CTS. Results were transmitted to the AE and BT schools.

CORRELATIONAL ANALYSES

Analyses conducted to assess the equivalence of the data collected by the three different methods are outlined here. The results of these analyses are presented in section III.

As mentioned previously, the primary objective of this study, from the TAEG viewpoint, was to determine the comparability, or equivalence, of the three methods of collecting training feedback information. For this, statistical comparisons of the data obtained were made to assess:

- the comparability of methods,
- the comparability of source (i.e., graduates versus supervisors),
- the comparability (or independence) of scales.

The various forms of equivalence were assessed through the use of the Pearson-Product moment correlation technique. This statistic yields a coefficient of correlation, r, which indicates numerically the degree of relationship between variables. Correlation coefficients may take on values ranging from 0 to plus or minus 1. High correlation coefficients indicate that distributions of ratings/scores are similar. Correlation does not address questions of similarity in magnitude.

In this study, high positive correlations between appropriate distributions of ratings/scores obtained by the different methods would support conclusions that the methods produce equivalent data. Similarly, high positive correlations between ratings on the same scales but obtained from different classes of respondents would support conclusions that the respondents judge the variable of interest similarly (i.e., sources of data are equivalent). High positive correlations between ratings taken from different scales might indicate dependent relationships; e.g., a high correlation between frequency and proficiency could mean that those tasks on which individuals are rated most proficient are those which are performed most frequently.

However, examination of the data is required to determine reasons for high correlations and to assess implications. If, for example, a high correlation is obtained between two sets of scores or ratings because they are both low, this would provide a signal that curriculum revision action should be considered. On the other hand, if the high correlation is due to the fact that both sets of ratings are high, this would indicate that training, for example, is adequate.

Table 3 shows the specific comparisons planned for evaluating the equivalence of data obtained by the three methods and from different sources. The "Xs" in each row of the table indicate that the two sets of ratings were directly correlated. The "Is" in each row indicate that the relationship was inferred from independent correlations of the two variables with a third variable. The logic underlying the inferred correlations is this: if scale A correlates with scale B, and scale B correlates with scale C, it is possible to conclude that A correlates with C. This conclusion is, of course, not based on a numerical value for the correlation. This subjective analysis is far from ideal, but it does permit an estimate of an otherwise indeterminate relationship. These inferred correlations were necessitated by methodological and/or statistical considerations which precluded the direct correlation of those variables. Two of these inferred correlations were the consequence of different samples being employed in the rating process. The third inferred correlation reflects the fact that no graduate received both the JKT and Q which makes a direct correlation impossible. An attempt was made in the BT sample to correct the latter deficiency but the Q return rate was inadequate (N=5) to perform the necessary analysis.

Correlations of JKT scores were made only with results of the SI. These correlations considered only the individuals who were given the JKT and also interviewed. All other possible correlations, not specifically identified in table 3, bear on the issue of independence of scales.

The planned comparisons described can only accomplish a single purpose--the evaluation of the comparability of training feedback data obtained through the various methods. They do not address the issue of data relevance or validity. The determination of validity requires a clear statement of the intended purpose for collecting the data. Hence, validity is defined by the extent to which the stated purpose has been accomplished. In training, the intended purpose for collecting feedback data might include the assessment of deficiencies in meeting training objectives, an assessment of deficiencies in

TABLE 3. PLANNED COMPARISONS FOR DETERMINING EQUIVALENCE OF METHODS AND DATA SOURCES

METHOD EQUIVALENCE	JKT: KNOWLEDGE	SI: GRADUATE: FREQUENCY	SI: GRADUATE: TRAINING ADEQUACY	SI: SUPERVISOR: FREQUENCY	SI: SUPERVISOR: TRAINING ADEQUACY	SI: SUPERVISOR: PROFICIENCY	Q: GRADUATE: FREQUENCY	Q: GRADUATE: TRAINING ADEQUACY	Q: SUPERVISOR: FREQUENCY	Q: SUPERVISOR: TRAINING ADEQUACY	Q: SUPERVISOR: PROFICIENCY
DIRECT COMPARISONS		X	X				X				
Q vs. SI				X	X	X		X	X	X	X
JKT vs. SI	X		X		X	X					
	X										
	X										
INFERRED COMPARISONS	I										I
JKT vs. Q	I							I		I	
	I										
SOURCE EQUIVALENCE											
SI		X		X							
			X		X						
Q							X		X		
								X		X	

required fleet performance resulting from training inadequacies, or simply to establish the relevance of training objectives. The determination of data validity was beyond the scope of this study. The responsibility for the determination of validity of the task ratings rests with the schools and the fleet. They are the only sources which can assess the degree to which the data are valid and relevant to their objectives.

Additional correlations were calculated to compare subarea scores of the JKT with applicable mean ratings of training adequacy given by graduates and supervisors on the SIs and supervisor's ratings of their proficiency. The purpose of this secondary analysis was to determine if specific relationships existed between JKT scores and mean scale ratings in one or more of the subareas identified in table 2. This analysis provides the respective "A" schools with data they might find useful in identifying specific strengths/weaknesses in their training. Note that it is also a finer grained analysis of data equivalency which verifies and expands on the primary correlation analysis. To perform this secondary analysis, correlations were computed for each subarea shown in table 2. That is, correlation coefficients were computed between corresponding item values (i.e., number right) on the JKT and rating scales for each subarea. The number of paired items varied in each subarea reflecting the construction of the JKT and item matching process described earlier.

RESOURCE ACCOUNTING

In addition to evaluating the three feedback gathering methods with respect to the equivalency of the data, it was also desired that information be compiled for evaluating them with respect to resources required. This would provide further guidance regarding which method might be most efficiently used in future data collection efforts.

To compile this information, records were maintained of labor, travel, and material costs required to collect the feedback data. The intent was, considering all relevant factors, to determine a "cost per instrument." These figures of merit[5] then could be used to support decisions concerning selection of data gathering techniques.

During the study, records were kept of man-hours, postage and printing costs, travel costs, and other cost factors. Also, PTEP interviewers were requested to record (on sheets provided by TAEG) items such as time to coordinate data collection efforts and to administer the instruments and fleet and own resources used.

It should be clearly recognized that this pilot study required the expenditure of considerable time and effort to accomplish tasks which would

[5] As used in this report, a figure of merit is defined as a numerical quantity based on one or more characteristics under specified conditions and used for indicating comparative efficiency or effectiveness.

not be required for future training appraisal efforts. For example, time and/or travel costs were incurred to perform tasks such as:

- briefing and coordinating the project at various command levels and with organizations external to the NAVEDTRACOM (e.g., NPRDC),
- defining the roles of the various participants in the study as well as for coordination between the principals on study details/mechanics,
- analyzing the data obtained and preparing briefings/reports of results,
- sustaining administrative delays occurring in the study,
- preparing special data summaries for answering experimental questions about the methods rather than evaluative questions about school training.

Unfortunately, resource information in the fine grained form required for defining costs pertinent to the evaluating of a particular curriculum was not separately maintained from total costs. And it would be grossly misleading to report total costs of the PEP which was essentially a research study. Consequently, the resources reported in section IV are estimates of resources that would be required to collect data via the three methods. These estimates are based on the past experience of both the TAEG and the CTS for PTEP in using the methods. Wherever justified, actual cost figures are used as the basis for the estimates.

ADMINISTRATIVE AND TECHNICAL PROBLEMS

A third information pool desired for evaluating the three data collection methods concerned administrative and technical problems associated with their use. To develop this information, informal records and notes were kept during the project for known problems. Also, several problems surfaced during analysis and interpretation of data. Identification and discussion of these problems are given in section V.

SECTION III

RESULTS AND DISCUSSION OF DATA EQUIVALENCE ANALYSES

This section presents the results of the correlational analyses performed to assess the equivalence of data gathered by the three collection techniques. Results are presented separately for each rating. Discussion and interpretations of the results are also presented.

RESULTS

Analytical results pertinent to the issues of method and source equivalence are presented below. Technical data supporting conclusions based on these results are given in appendices C through F.

AE Q AND SI INTERSCALE CORRELATIONS. Table 4 presents the results of the correlation analyses performed on the AE Q and SI data to determine the degree of relationship (i.e., equivalence of data) between these techniques. The numerical designators in the rows and columns of the table are provided to simplify locating correlation values of interest. The first designator identifies a row, the second a column. For example, location 9,4 identifies the row/column intersection at which the correlation ($r = .77$) between supervisor's mean ratings of training adequacy given on the Q and supervisor's mean ratings of training adequacy given on the SI may be found. All correlations were computed for mean ratings obtained over the 124 job tasks of the Q and SI instruments.

Correlations bearing on the equivalence of the two methods are contained in area II of the table, specifically at the diagonal formed by locations 6,1; 7,2; 8,3; 9,4; and 10,5. Locations 1,3 and 2,4 (area I) and 6,8 and 7,9 (area III) relate to the equivalence of the source of data (i.e., graduates versus supervisors). The remaining correlations in the table assess the issue of equivalence (or independence) of scales between and within the Q and SI. The pattern of these correlations was different for the AE and BT ratings. These differences are discussed at the end of this section.

AE JKT AND SI CORRELATIONS. Two sets of correlations were computed for assessing overall relationships between JKT scores and ratings assigned by the SI method. For both, correlations between JKT results and SI ratings were low and nonsignificant.

Table 5 shows the correlations between JKT scores and SI ratings. These are for the 32 individuals given both the JKT and SI. The correlations shown are between the distributions of total number right (i.e., score for each individual) on the JKT and mean ratings (over the 124 job tasks) for a given scale of the SI.

TABLE 4. INTERCORRELATIONS OF MEAN SCALE RATINGS GIVEN BY AEs ON THE Q AND SI INSTRUMENTS BY GRADUATES (G) AND SUPERVISORS (S)*

	SI Responses					Q Responses				
	(Gf)1	(Gta)2	(Sf)3	(Sta)4	(Sp)5	(Gf)6	(Gta)7	(Sf)8	(Sta)9	(Sp)10
SI Responses			I							
G-Frequency Scale (Gf) 1		.04	.93	-.16	.03					
G-Training Adeq. (Gta) 2			.02	.57	.62					
S-Frequency Scale (Sf) 3				-.22	-.07					
S-Training Adequacy (Sta) 4					.52					
S-Proficiency (Sp) 5										
Q Responses			II					III		
G-Frequency Scale (Gf) 6	.93	-.03	.92	-.27	-.08		.-11	.97	-.28	.23
G-Training Adeq. (Gta) 7	-.04	.71	-.11	.57	.66			-.07	.76	.61
S-Frequency Scale (Sf) 8	.93	.01	.95	-.22	-.04				-.24	.29
S-Training Adequacy (Sta) 9	-.17	.67	-.25	.77	.73					.69
S-Proficiency (Sp) 10	.32	.59	.23	.58	.55					

* Values above r = .18 significant at the .05 level for df = 122

TABLE 5. CORRELATIONS BETWEEN AE GRADUATES JKT SCORES AND RATINGS OF TRAINING ADEQUACY AND PROFICIENCY ON THE STRUCTURED INTERVIEW*

	SI SCALE		
	G-Training Adeq.	S-Training Adeq.	S-Proficiency
Score On JKT	-.08	.18	-.10

* Value of r needed for significance at the .05 level for 30 df is .35

Table 6 shows the correlations obtained between the group mean number correct on the JKT and mean SI rating values for the 47 matched items, where the means are based on a sample of 32 respondents.

TABLE 6. CORRELATIONS BETWEEN AE JKT SCORES AND SI MEAN SCALE VALUES ON CORRESPONDING ITEMS*

	SI SCALE		
	G-Training Adeq.	S-Training Adeq.	S-Proficiency
Mean No. Right On JKT Items	.14	.07	.03

* Value of r needed for significance at the .05 level for 45 df is .29

Table 3 (section II) shows that the equivalence of JKT and Q data was to be assessed by inference. Note that the inferred comparisons, however, require significant correlations between the JKT and the various scales of the SI for the inferential logic to be carried out. The nonsignificant values between the JKT and the scales of the SI preclude any such attempt.

AE Subarea Correlations. Table 7 shows the correlations between JKT subarea scores earned by the 32 graduates and SI subarea mean scale values on corresponding items. The table shows the letter designation given by PTEP to the subarea, the name of the subarea, and the number of pairs on which each correlation was based. The values of "r" obtained for the correlations between the JKT subarea scores and mean values for the corresponding task ratings for the Graduate Training Adequacy (Gta) scale, the Supervisor Training Adequacy (Sta) scale, and the Supervisor Proficiency (Sp) scale are also shown. The final column of the table lists the value of "r" needed to conclude that there is a significant relationship between the two sets of variables. A correlation coefficient was not computed for subarea I since this statistical procedure cannot be performed on only one pair of scores.

TABLE 7. CORRELATION VALUES OBTAINED BETWEEN JKT SUBAREA SCORES AND SI SUBAREA MEAN SCALE RATINGS (AE)

SUBAREA Letter	SUBAREA Name	No. of Pairs	SI SCALE Gta	SI SCALE Sta	SI SCALE Sp	r Needed for Sig.
C	Wiring	9	.12	-.03	-.15	.67
D	Basic Troubleshooting	7	.16	.45	.47	.75
E	Supply and Administration	3	.994	.97	.98	.997
F	Aircraft Instruments	12	.19	-.04	.02	.58
G	Engine Performance/Warning Systems	6	-.42	-.38	-.13	.81
H	Electrical/Electronic Systems	3	.65	1.00	.993	.997
I	Auto Pilot Systems	1	-	-	-	-
J	Scheduled Maintenance	7	-.08	-.72	-.42	.75

Only the correlation between subarea H and the Supervisor Training Adequacy scale could be considered significant. However, extreme caution should be used in evaluating the practical significance of the data in the table since they are subject to sampling error associated with small samples.

BT Q AND SI INTERSCALE CORRELATIONS. Table 8 shows the intercorrelations of the scales of the Q and SI instruments for the BT rating. The table should be read in the same way as table 4. All of the obtained correlation coefficients were significant beyond the .05 level.

BT JKT AND SI CORRELATIONS. Similar to the findings for the AE rating, correlations between JKT scores and the rating scales of the SI were low and not significant. This again precluded attempts at establishing inferred relationships between the JKT and Q.

Table 9 shows the correlations between the JKT scores and appropriate SI scales for the 51 individuals given both the JKT and the SI. These correlations are between the individuals' scores on the JKT and overall mean ratings given by respondents to the 124 job tasks on a given scale.

TABLE 8. INTERCORRELATIONS OF MEAN SCALE RATINGS GIVEN BY BTs ON THE Q AND SI INSTRUMENTS BY GRADUATES (G) AND SUPERVISORS (S)*

	SI Responses					Q Responses				
	(Gf)1	(Gta)2	(Sf)3	(Sta)4	(Sp)5	(Gf)6	(Gta)7	(Sf)8	(Sta)9	(Sp)10
SI Responses			I							
G-Frequency Scale (Gf) 1		.48	.95	.51	.68					
G-Training Adequacy (Gta) 2			.48	.63	.57					
S-Frequency Scale (Sf) 3				.55	.71					
S-Training Adequacy (Sta) 4					.86					
S-Proficiency (Sp) 5										
Q Responses			II				III			
G-Frequency Scale (Gf) 6	.94	.44	.93	.48	.66		.50	.94	.55	.72
G-Training Adequacy (Gta) 7	.58	.80	.55	.70	.64			.49	.69	.58
S-Frequency (Sf) 8	.89	.41	.92	.45	.63				.59	.75
S-Training Adequacy (Sta) 9	.54	.55	.60	.68	.63					.81
S-Proficiency (Sp) 10	.69	.47	.73	.55	.71					

* Values above r = .18 significant at the .05 level for df = 122

TABLE 9. CORRELATIONS BETWEEN BT GRADUATES JKT SCORES AND RATINGS OF TRAINING ADEQUACY AND PROFICIENCY ON THE STRUCTURED INTERVIEW*

	SI SCALE		
	G-Training Adeq.	S-Training Adeq.	S-Proficiency
Score on JKT	.23	-.20	.04

*Value of r needed for significance at the .05 level for 49 df is .276

Table 10 shows the correlations obtained between the mean number right on JKT items and mean SI rating value for the 21 pairs of items where the means are based on a sample of 51 individuals. Again, none of the correlations of JKT with scale values was significant.

TABLE 10. CORRELATIONS BETWEEN BT JKT SCORES AND SI MEAN SCALE VALUES ON CORRESPONDING ITEMS*

	SI SCALE		
	G-Training Adeq.	S-Training Adeq.	S-Proficiency
Mean No. Right on JKT Items	.23	.25	.38

*Value of r needed for significance at the .05 level for 19 df is .43

BT Subarea Correlations. Table 11 shows the correlations obtained between SI mean scale values for subareas and mean JKT scores on corresponding items. None were significant. For the BT rating, only 21 items were matched. This is sufficient for conclusions regarding the overall relationship between JKT scores and scale values. However, when these 21 items are apportioned among the eight subareas of the BT rating, the number of items in each subarea, in all but three instances, was too few to permit statistical comparisons. By inference, however, it can be concluded that (because of the nonsignificant overall correlation values and because of the nonsignificant AE subarea scores) the JKT subarea scores are not related to the SI answers of the group.

TABLE 11. CORRELATION VALUES OBTAINED BETWEEN JKT SUBAREA SCORES AND SI SUBAREA MEAN SCALE RATINGS (BT)

SUBAREA			SI SCALE			
Letter	Name	No. of Pairs	Gta	Sta	Sp	r Needed for Sig.
C	Boiler Operation	10	.45	.38	.50	.63
D	Boiler Maintenance	3	-.79	.80	.70	.997
E	Boiler Cleaning	3	.53	.73	.71	.997

INTERPRETATION AND DISCUSSION OF RESULTS

The following paragraphs are devoted to an interpretation and discussion of the results of the data equivalency analyses presented above. The issues of comparability of the different data collection methods and of sources of data (i.e., graduates versus supervisors) are discussed. These are given separately for each rating. The issue of comparability (or independence) of the scales of the Q and SI instruments while secondary to this study--it has implications for instrument design--is briefly noted in the final section which discusses differences observed in the data as a function of rating.

Q VERSUS SI METHOD (AE). The comparability of results obtained from the Q and SI instruments is discussed below for each scale of these instruments.

Equivalence of Frequency Data (AE). Correlations r 6,1 (.93) and r 8,3 (.95) of table 4 indicate a significant relationship between the ratings of task performance frequency obtained on the Q and the SI. Correlations r 1,3 (.93) and r 6,8 (.97) of the same table indicate that supervisors and graduates produce comparable frequency data on both the Q and SI.

It is therefore reasonable to conclude that equivalent frequency data can be obtained from either the Q or the SI and from either graduates or supervisors. However, it should be noted that in this study the SI was actually an orally administered Q. In the event that an interview technique other than that employed here is utilized in gathering feedback information, it will be necessary to establish the nature of the relationship between the data obtained through that technique and questionnaires.

Equivalence of Training Adequacy Data (AE). Correlations r 7,2 (.71) and r 9,4 (.77) indicate a similarity in training adequacy data obtained via the Q and SI. However, this relationship is not as well established or as straightforward as that observed in the frequency data. This may be at least partially attributed to the unreliability associated with highly complex judgments of subjective variables such as quality or value. On the other hand, judgments of event frequency are less complex depending primarily upon the accuracy of recall. Nevertheless, the trends in the data support the conclusion that either the Q or SI may be used to collect data on training adequacy.

Correlations r 2,4 (.57) and r 7,9 (.76) are equivocal with regard to whether training adequacy data obtained from graduates or supervisors are equivalent. The discrepancy in the absolute values of these correlations suggests a cautious approach to their interpretation. But, the trends in the data support the conclusion that supervisors or graduates can be considered equivalent sources of training adequacy data.

Equivalence of Proficiency Data (AE). The assessment of proficiency of graduates was, of course, limited to data obtained from supervisors. Correlation r 10,5 (.55) of table 4 statistically supports the conclusion that assessment of proficiency can be accomplished by using either the Q or SI. However, the absolute value of the correlation is of marginal utility in helping to decide whether to use the Q or SI in a practical setting. The coefficient of determination for this correlation (obtained by squaring the correlation coefficient) is approximately .30 indicating that 70 percent of the variance cannot be attributed to a systematic relationship between the two variables.

Interpretation of the correlations between the proficiency scale and other scales must also be made in the light of certain other technical considerations. First, supervisors in the Q group rated the proficiency of a different group of graduates than those in the SI group. To the extent that graduates in these two groups actually differed in proficiency, one would expect a less than perfect relationship between the ratings obtained from the two different instruments. Here, assessment of the equivalence of the two methods would have been better done by having supervisors rate the same graduates twice--once under each technique. This would provide a type of reliability check across the instruments and would provide information bearing more directly on their equivalence.

The issue of equivalence, however, is somewhat secondary in the case of proficiency ratings. Ratings of proficiency can be made equally well under a variety of techniques which call for the same types of judgments. More important to performance rating than the technique used to collect them is the issue of reliability, or consistency, with which such ratings are made. This is especially relevant when the ratings are intended to serve as a validity criterion, as they were in this study. The proficiency scale used here was not intended to provide any direct information bearing on the quality of "A" school training. It was intended, rather, to provide a criterion for validating the JKT as discussed below.

Hardly surprising is the fact that both graduate and supervisor ratings of adequacy correlate significantly with supervisor ratings of proficiency on the Q ($r_{2,5} = .62$; $r_{4,5} = .52$) and the SI ($r_{7,10} = .61$; $r_{9,10} = .69$), respectively. Graduate and supervisor ratings on the Q also correlate significantly with supervisor ratings of proficiency on the SI ($r_{7,5} = .66$; $r_{9,5} = .73$). Similarly, supervisor ratings of proficiency on the Q are positively related to graduate and supervisor ratings of adequacy on the SI ($r_{10,2} = .59$; $r_{10,4} = .58$). Overall, these findings show a statistically significant, although moderate, relationship between ratings of adequacy and proficiency regardless of whether the data are obtained via the Q or SI.

One basis for this set of findings is the apparent logical relationship between training adequacy and proficiency of graduates. That is, if training is considered adequate (as it was by both graduates and supervisors) then graduates from such a program would in all likelihood be considered proficient. This conclusion could be reached as the result of a real relationship between adequacy and proficiency or from a logical rating error in which such a relationship is merely assumed. However, analysis of the data leads to the conclusion that redundant information is being assessed by the two scales and, therefore, one of these scales can be eliminated in future efforts with little loss of information. Since the concern of these measures is with the evaluation of training, the proficiency scale probably should be eliminated since it is oriented toward performance which is best measured by more direct means. In addition, ratings of proficiency tended to be higher than ratings of adequacy which could produce an over estimation of training quality if employed in training evaluation.

JKT VERSUS SI (AE). No significant correlations were found between the total number right on the JKT and mean SI ratings on the graduate training adequacy scale (-.08), supervisors training adequacy scale (.18), or supervisor ratings of proficiency (-.10) (see table 5). Similar results were obtained for matched item correlations (table 6) on the same variables (r's = .14, .07, and .03), respectively. In simple terms, this means that the JKT and the SI measure different things. They do not provide equivalent data. It should be noted that the JKT was designed to assess factors relating to job performance while the SI and Q were designed to assess factors relating to training. Since the methods involved different purposes and were not compared in terms of their relationship to a common performance standard(s), the determination of a "best" method of obtaining training feedback data seems inappropriate. At present, such a determination hinges on an evaluation of resource requirements and administrative and technical problems associated with the various methods. However, it must be pointed out that the JKT scores did not correlate significantly with the proficiency ratings criterion developed for assessing validity of the JKT. It is likely that supervisors did consider job knowledge in rating the overall proficiency of graduates to perform job tasks. To the extent that they did, the JKT was insensitive to measuring it.

AE Subarea Correlations. Since there was no significant overall relationship between JKT scores and ratings of training adequacy or proficiency on the SI,

an additional analysis was conducted to determine if specific relationships existed in various subareas of those measures (see table 7). Examination of those data reveals a pattern of nonsignificant correlations which verifies the conclusion reached with regard to the overall analysis. That is, the JKT and SI appear to be measuring different things. Even those correlations that were significant (subarea H) or approaching significance (subarea E) were based on such small samples (N = 3) that these specific correlation values could easily be attributed to chance and should, therefore, be considered of little practical significance.

JKT VERSUS Q. As noted previously, determination of the relationship between data obtained on the JKT and from Qs depended upon the demonstration of significant correlations between SI ratings and JKT scores. Since these correlations were not significant, the inferences necessary to support conclusions regarding the equivalency of JKTs and Qs could not be made. However, because of the substantial correlations between Qs and SIs plus the low correlations between SIs and JKTs, it is a reasonably safe hypothesis that no significant relationships would be found between Q and JKT data.

METHOD COMPARABILITY (BT). The major results obtained with the BTs are so nearly identical to those obtained with the AEs that only a summary of conclusions regarding the BTs will be presented. The same cautions indicated in interpretation of the AE data, general conclusions reached, and other considerations discussed are appropriate here unless otherwise indicated.

With respect to equivalency of the Q and SI methods and data sources, the following conclusions are possible from table 8.

- Equivalent frequency data can be obtained from either the Q or the SI ($r_{6,1} = .94$; $r_{8,3} = .92$)
- Supervisors and graduates produce comparable frequency data on both the Q and SI ($r_{1,3} = .95$; $r_{6,8} = .94$)
- Equivalent training adequacy data can be obtained from either the Q or the SI ($r_{7,2} = .08$; $r_{9,4} = .68$)
- Equivalent training adequacy data can be obtained from either graduates or supervisors ($r_{2,4} = .63$; $r_{7,9} = .69$)
- Equivalent proficiency data can be obtained from either the Q or SI ($r_{10,5} = .71$).

Similar to the AE findings, all combinations of correlations between Q and SI training adequacy and proficiency scales were positive and significant (see table 8, locations 2,5; 4,5; 7,5; 7,10; 9,5; 9,10; and 2,10).

With respect to the equivalency of JKT and SI methods, the following conclusions are possible.

- The JKT measures different things than do the training adequacy (coefficients of .23, -.20, .23, and .25) and proficiency scales of the SI (coefficients of .04 and .38) (see tables 9 and 10).

- There were no significant relationships between JKT subarea scores and SI subarea mean scale ratings (see table 11).

No firm conclusions can be made concerning the equivalence of Q and JKT data.

AE VERSUS BT DATA. Tables 4 and 8 show significant differences in the values of certain of the Q/SI interscale correlations obtained for the AE and BT respondents. The major differences occur in those correlations which were not part of the comparisons planned to assess method or source equivalence (see table 3). Even though they are inconsequential to the determination of method equivalence, they are mentioned here since they bear on the issue of scale independence which is important to instrument design. Accordingly, such relationships should be considered in future scale design/selection efforts.

The most prominent feature of the data is that all interscale correlation coefficients were significant for BTs (table 8) while a high proportion of the corresponding correlations for the AEs (table 4) were not significant. Appendix E lists the specific correlations and values in question. Relationships which were expected to be significant; i.e., those between ratings of training adequacy and proficiency, which were discussed earlier are not included. The average value of these correlations (n = 24) was -.05 for the AEs and .57 for the BTs. Thus, the scales appear to be independent for the AEs but not for the BTs. This finding precludes any firm conclusions concerning scale independence.

OVERALL SCALE MEANS

As mentioned previously, correlation techniques are not sensitive to differences in magnitude between the values of the variables in distributions compared by this statistical technique. To support conclusions of equivalence of data sets reached by correlational analysis, it is also desirable that the absolute values of the variables in question be examined for gross discrepancies. This was done by computing the overall scale means obtained on the Q and SI instruments from graduates and supervisors. These data are reported in appendix F.

SECTION IV

RESOURCE REQUIREMENTS AND DISCUSSION

This section presents and discusses resource requirements associated with the collection of data via the Q, SI, and JKT. Actual cost figures are used where appropriate. In compiling the resource information, the intent was to develop and present data that would be useful for direct estimation of the costs associated with future use of each method. The data summaries permit establishment of a relative cost per instrument based on the number of usable returns. These figures of merit can be used to estimate and compare anticipated evaluation costs for different situations.

RESOURCE DATA

Resources expended to collect training feedback data are presented separately for each rating. Costs incurred for the pilot study which would not be factors in future evaluation efforts (e.g., time and travel to brief and coordinate the project, to define the roles of the various participants, to prepare special analyses, and to prepare the present report) are not included in the data summarized below.

AE. Table 12 shows a summary of costs for the AE portion of the study for labor, travel, and material. All costs are rounded to the nearest $10.

TABLE 12. AVIATION ELECTRICIAN'S MATE COST DATA

RESOURCE CATEGORY	DATA COLLECTION METHOD					
	Q		SI		JKT	
	MAN-HOUR	COST	MAN-HOUR	COST	MAN-HOUR	COST
LABOR						
Development	135	$1,040	165	$1,300	860	$7,870
Instrument Administration (actual)						
a. Interviewer/ Monitor	33	355	115	1,025	48	483
b. Fleet Respondent	197	1,075	115	615	272	1,167
Analysis	475	4,450	470	4,430	400	4,030
TRAVEL		$1,050		$1,350		$1,200
MATERIAL						
Printing/Postage		$ 750		$ 180		$ 90
Computer Processing		600		600		600
Reporting		130		130		30
TOTALS	840	$9,450	865	$9,630	1,580	$15,470

Development labor costs are based on time and salary of the actual developers (GS-11/5 for the Q and SI and E-9 for the JKT). The labor costs for typing support (12 hours) and an E-6 subject matter expert's assistance (8 hours) were added to the development costs for each method. It was estimated that 80 man-hours were required to develop a survey instrument (Q/SI) which contained 124 task statements.

Since the same instrument was used for both the Q and SI, the same developmental labor costs were used in computing total costs as if those instruments were independently developed. Preadministration or pretest costs are also included in developmental costs. The different dollar values in table 12 for development of the Q and SI instruments are due to different pretest procedures. SIs require a one-on-one condition for pretest whereas Qs can be pretested effectively in a group setting. The JKT was pretested on three groups of 12 persons each in the Norfolk area.

Development costs for JKT are substantially higher. The guidelines provided by CTS for PTEP personnel for estimating test development costs identify a requirement for the development of an initial test bank of 200 questions. This pool is subsequently reduced to 100 items by rational and statistical analysis. Multiple choice questions average 2 hours each for development, while a skill exercise usually requires 160 hours to develop. The AE JKT contained 100 multiple choice questions and two skill exercises.

The costs incurred in the administration of the instruments are reported in the second row in the table. Forty-seven questionnaires were received from graduates and 55 from supervisors. Average time to complete the Q was 1.9 hours. Structured interviews were conducted with 56 graduates and 32 supervisors. Average time to complete the SI was 1.1 hours. The JKT was administered to 85 graduates. Personnel time includes the individual(s) administering the Q, SI, or JKT and the individual(s) completing the instrument. A salary schedule for an E-3 was used for the graduate, an E-6 for the supervisor, and an E-8 for the interviewer/test administrator (the average grade level for interviewers was E-8). Salary rates were obtained from the Navy Billet Cost Model (1977-1978) and the General Schedule Salary Rate (1977-1978). The 18 hours required to coordinate interviews/tests were also included. Most arrangements were made by E-9s at PTEP. These were made with fleet personnel ranging from Chief Petty Officers to Lieutenant Commanders. One 2-hour briefing was required for obtaining approval to distribute the Q.

The labor used to analyze the data collected includes the effort for development of data reduction procedures, processing routines, and formats. Work done by programmers and encoders was also incorporated into the summary of analysis costs. The estimate is inflated for the AEs due to the considerable effort required to recode data from the Q/SI booklets onto OPSCAN booklets for machine processing at NOTAP. This cost would not be repeated with the use of OPSCAN answer sheets. The JKT data were recorded on OPSCAN sheets. An estimated 1 hour was required to process these answer sheets. Two resources were applied equally across the methods: programming time and report writing. The development and debugging of a routine program was estimated at 80 hours. Finally, 2 hours

per page were projected for report writing.

Travel costs shown for the Q/SI effort resulted from two trips to the AE school at Memphis and one trip to Norfolk. These were for instrument development, pretesting of the instruments, and data analysis. Additional travel was required for PTEP personnel to conduct SIs and administer JKTs in Jacksonville, FL., and Norfolk, VA. The PTEP staff also made two trips to Memphis: one for JKT development and one during the analysis phase.

The last category, "Material," summarizes the costs of printing 400 Qs and 100 SIs (each 26 pages long). PTEP costs were for xeroxing 30 JKT booklets of 50 pages each and for the printed OPSCAN answer sheets. The cost of envelopes and postage used to send out and receive completed Q forms is also included. Accurate records of time were not maintained since required data processing was done on an as-convenient and time-available basis. However, 30 minutes on a large computer at a cost of $20 per minute would be a reasonable estimate of processing time requirements for each method.

BT. Table 13 summarizes the resources required for the BT phase of the project. Since the BT plan was highly similar to the AE plan, the resource costs are accordingly similar. Labor cost differences result from differences in the number of personnel pretested, the number surveyed, and the number of returns.

TABLE 13. BOILER TECHNICIAN'S COST DATA

RESOURCE CATEGORY	DATA COLLECTION METHOD					
	Q		SI		JKT	
	MAN-HOUR	COST	MAN-HOUR	COST	MAN-HOUR	COST
LABOR						
Development	165	$1,370	180	$1,520	845	$7,830
Instrument Administration (actual)						
a. Interviewer/ Monitor	33	355	113	1,002	72	725
b. Fleet Respondent	92	525	112	638	638	2,545
Analysis	470	4,430	475	4,450	415	4,030
TRAVEL		$1,250		$1,280		$ 830
MATERIAL						
Printing/Postage		$ 550		$ 50		$ 100
Computer Processing		600		600		600
Report		100		100		30
TOTALS	760	$9,180	880	$9,640	1,970	$16,690

BT Q/SI task statement development required more time than AE development because of the need to familiarize developers with the subject matter. Pretest of the Q/SI involved 13 BTs. Again, the one-on-one interview results in a higher cost for the SI than for the Q. The same cost factors apply to the development of the BT JKT as the AE JKT with the exception of the number of persons pretested (30).

Administration costs are based on actual returns. There were 90 completed Q forms from graduates and supervisors and 102 completed SIs. Average time to complete both instruments was 1 hour. Two hundred and nine JKTs were completed. The rationale described for establishing AE administrative costs also applies to the BT administrative costs.

Analysis and support services were provided by the same level personnel as those in the AE study. Cost differences are due to the number of returns and encoding procedures. The Q/SI data were keypunched directly into the WANG Programmable Calculator. This direct data entry took less time than recoding booklets to be opscanned.

Travel costs for the BT effort were for trips to the school at Great Lakes. SI/JKT data were collected in the Charleston and Norfolk areas. The remaining costs were the same as for AEs.

Cost factors for material are the same as shown in the AE table. However, the number of BTs surveyed was higher and no costs for NOTAP OPSCAN booklets were incurred.

COST PER INSTRUMENT. Cost per instrument incurred during the PEP and an example of an application (nondevelopmental) are discussed below.

PEP Costs. The bottom line figures of tables 12 and 13 can be divided by the number of usable returns to provide an average cost (figure of merit) per usable return associated with the use of each method. Table 14 provides a summary of actual costs per usable return and average (projected) costs based on 100 returns.

TABLE 14. ACTUAL AND PROJECTED COST PER INSTRUMENT

ACTUAL COSTS	Q	SI	JKT
AE	$ 92.65	$109.43	$182.00
BT	102.00	94.51	79.86
PROJECTED COSTS			
AE	$ 94.50	$ 96.30	$154.70
BT	91.80	96.40	166.90

The large discrepancy between AE and BT actual cost for the JKT reflects large differences in the number of usable returns obtained in the two efforts. There were 209 usable JKT returns for the BTs and only 85 for the AEs. This reflects an amortization of development costs as the number of usable returns increases. Such cost decreases, however, will be offset by increased administrative costs for the SI because of the one-on-one interview conditions.

The figures of merit discussed above can be used as indices of the relative costs of feedback data collection via Qs versus SIs versus JKTs. Given the same number of returns per instrument, the Q emerges as the least expensive method and the JKT as the most expensive. The absolute dollar values presented in the table reflect costs incurred during the PEP. Under different sets of conditions (e.g., shorter instruments, less administrative time, higher Q return rates, larger numbers of respondents) such as would occur in "routine" training evaluation situations, the absolute costs associated with each instrument will decrease. This is demonstrated by the discussion given below which generates a cost per instrument based on the collection of data from 300 respondents per instrument. The same cost factors were applied as in the previous tables.

Estimated Appraisal Costs. Table 15 develops and presents a "per instrument cost" for a representative training appraisal effort. Costs are rounded to the nearest 10 cents. The following assumptions were used in generating the data for this example:

- A detailed data base consisting of information about the course curriculum, specialty job tasks, and skill requirements is available for use in instrument development.
- Each instrument contains 100 task statements/questions.
- The JKT contains two skill exercises.
- The salary level of administrator is GS-11/5.
- The salary level of respondents is E-3.
- The time to complete the Q or SI instrument is 1 hour.
- The time to complete the JKT with skill exercises is 2 hours.
- There are 30 individuals in each preadministration group.
- Costs are not amortized over time.
- Three hundred graduates are assessed by each method.

Row 1 of table 15 shows the prorated development costs for each instrument. The paragraphs below explain the bases of these estimates. Based on PEP data, it would take 40 minutes to develop each Q task statement. For an instrument of 100 items the development cost without any pretest is $643. An additional $100 is required for typing and SME review. Pretesting the instrument on 30 E-3s (in three sessions of 10 individuals each) costs $145. Thus, total development cost

TABLE 15. COMPARATIVE COSTS OF DATA COLLECTION

COST FACTORS	INSTRUMENT		
	Q	SI	JKT
1. Prorated Development Cost	$3.00	$3.80	$24.00
2. Salary of Instrument Administrator (1 hour)		9.65	1.00 (prorated)
3. Salary of Respondent	3.90 (1 hr)	3.90 (1 hr)	7.80 (2 hrs)
4. Administrative/Clerical Salary (per minute)	.80 (5 mins)	1.60 (10 mins)	.80 (5 mins)
5. Postage (3rd class out/1st class return)	.90	-	-
6. Travel (1,000 miles) (Obtain data 300 people)	-	9.80 6/day	4.00 2/day
7. Total Cost/Application	8.60	28.75	37.60

rounded to the nearest decile is $900 for the questionnaire. Prorating the development costs over 300 individuals yields an average cost of $3 for each Q.

Similarly, for an SI instrument of 100 items, the development cost without pretest is $643 plus $100 for typing and SME review. The pretest is more expensive ($405) due to the one-on-one interview method. Total development cost for the SI is $1,150 in this example and the "per instrument cost" is $3.80.

The cost guidelines provided by PTEP indicate that the development costs for a 100 item JKT would total $3,860. This includes the development of the initial item pool. The development of two skill exercises would increase JKT costs to $6,950. An additional $100 is required for typing and SME review. A pretest given in three groups of 10 each will cost $145. Thus, the total development costs for the JKT amount to $7,190 and the "per instrument cost" is $24.

Costs associated with collecting data by each method are apportioned among rows 2 through 6 of table 15. Row 2 shows salaries associated with instrument administration. The salary for the JKT Administrator is prorated to reflect group testing (i.e., 10 per session). The third row shows salary for E-3 respondents. Miscellaneous administrative and clerical time associated with project coordination and record keeping functions are shown in row 4. Postage costs for the Q, shown in row 5, are based on an estimated 10 page long OPSCAN booklet mailed third class and returned first class.

For the example being discussed, a travel distance of 1,000 miles was assumed for both the SI and JKT. Estimated time to collect data from 300 graduates was used to determine length of TDY. Six SIs and two JKT groups of 10 each per day were judged to be reasonable estimates of the number of instruments that could be administered in a single day. Airfare and rental car costs (at $20 per day) are included in the amounts in row 6. Row 7 shows the estimated per instrument cost for each of the three methods. Note that if skill exercises are not included with the JKT, instrument development costs, respondents salaries, and travel costs would be reduced. This would result in a per instrument cost of $21.38.

DISCUSSION

On the basis of the cost information provided above, it is clear that the Q is the least expensive method for collecting the type of feedback information gathered during the PEP. However, in the final choice of method, a number of factors other than cost must also be considered.

A previously published TAEG report (Hall, Rankin, and Aagard, 1976) provided a conceptual framework for planning and conducting training effectiveness assessments. It also provided discussions of the strengths, weaknesses, and inherent limitations of the various data gathering techniques and their appropriateness for use in different types of training appraisal situations. As noted in the report, the primary determinant of choice of data collection method (and the specific design of instruments/procedures used for data collection) should be the information that is required from an appraisal effort. While costs of data collection are important they must be considered in conjunction with the issue of whether a particular method will yield the proper kind and level of information

needed for decisions about training. Obviously, training feedback information which does not meet this requirement is valueless regardless of how inexpensively it can be collected.

The three methods evaluated in this study--questionnaires, structured interviews, and job knowledge tests--when properly employed will yield distinctly different kinds of information about training effects. It must be noted that the three data collection techniques evaluated in the PEP are not fully representative of their respective classes. They are simply examples of these classes. Questionnaires of the type used here are best suited for obtaining gross information about how adequate a particular training course is in preparing individuals for fleet jobs. If decisionmaking needs can be met by this type of information, it can be collected relatively inexpensively. This technique by itself, however, yields relatively little information concerning what specifically to change or retain in a training course. Thus, if the information needed from an appraisal effort concerns the nature of specific inadequacies (or strengths) of a course, then a much more detailed questionnaire will be required at considerably higher developmental costs. In this instance a carefully considered SI procedure could probably better be employed to obtain the necessary information at less cost and with less overall interference to fleet activities. The conventional SI procedure (not the type used in this study) is uniquely suited to in-depth probing to clarify and pinpoint specific issues. In the hands of a skilled interviewer, this technique will likely produce much more information per administration than a corresponding questionnaire. This feature means that fewer people need to be surveyed to obtain useful information.

As discussed in the preceding section, the job knowledge test is designed for a different purpose than the assessment of training course deficiencies. These tests, as currently designed and employed, are intended for use as tools in assessing the possession by job incumbents of the factual knowledge needed to support equipment operation/maintenance. The submarine community's acceptance and use of these tests attest to the fact that they are well suited for this purpose. In the opinion of TAEG, however, procedural and technical modifications to this testing concept are needed before they can be recommended for routine widespread use in "external" training appraisal programs. This testing concept is ideally suited for use at the schoolhouse level as a means of determining whether a course, in fact, taught necessary job knowledges--and its use is recommended at that level for that purpose.

In summary, appropriately designed questionnaires can be used relatively inexpensively (given favorable return rates) as screening devices to assess the overall adequacy of training for particular school-trained job tasks. A special problem may emerge, however, with respect to "C" school graduates as there may be too few judges onboard a particular ship who can assess the "adequacy" of such training. Here, JKTs given to course graduates may be the only viable way (short of performance testing) of assessing their achievement. Structured interviews, conducted with a selected number of individuals (preferably supervisors), can be effectively used for determining specific training deficiencies, probably at an acceptable cost because of the intrinsic strength of this technique in eliciting specific information. Given the current design purposes and testing concepts underlying JKTs, they are not recommended for routine external training appraisal purposes, but this testing concept could be extremely

beneficial at the school level. Finally, it should also be noted that there are other techniques available for training appraisal purposes than those addressed in the PEP (Hall, et al., 1976). Their use may be more suitable to filling specific information needs and less costly than either questionnaires, structured interviews, or job knowledge tests.

SECTION V

DISCUSSION OF ADMINISTRATIVE/TECHNICAL PROBLEMS

A number of problems of an administrative and/or technical nature arose during the PEP study. For the most part, these "problems" were resolved as they arose. In the process, however, a number of valuable lessons were learned. Since these have direct applicability to the conduct of future evaluation efforts, they are discussed briefly below. Where appropriate, suggestions are made for avoiding the same kinds of problems in the future.

FLEET COORDINATION

Responsibility for coordinating the PEP with the fleet was assigned to the CTS for PTEP acting in a "test" capacity as a Fleet Feedback Data Collection Group (FFDCG). This group briefed appropriate Atlantic fleet representatives on the purposes of the study and the cooperation desired from the fleet.

Some difficulties were encountered initially in securing Atlantic fleet approval for, and endorsement of, the PEP study. These difficulties were subsequently resolved through the personal intervention of the Commander, Training Command, U.S. Atlantic Fleet (COMTRALANT). However, the beginning of data collection efforts in this fleet were significantly delayed. And, because of this delay, no attempt was made to obtain Pacific Fleet participation as this process might well have resulted in still further delays to the PEP schedule. Once upper level authorization was obtained, cooperation at the individual unit level was excellent in all cases.

These experiences are related simply to point out the need for establishing mechanisms for obtaining timely fleet authorization/cooperation for gathering training feedback from fleet sources. Such mechanisms could be established to operate on either ad hoc or more permanent bases to avoid lengthy schedule delays in future training appraisal efforts.

STUDY SAMPLES

In this study, individual graduates served as the focal points of evaluation. Names and current duty station assignments/addresses of recent graduates of AE and BT courses were obtained from the respective schools by PTEP. These listings were used to select the individuals to be contacted during the study.

It was found that in many cases the duty station addresses supplied were incorrect. In future evaluation efforts, it may be desirable to draw random samples from graduate populations or to compose samples from listings such that they are representative of, for example, various classes of ships or types of aircraft. For this, accurate knowledge of graduate locations will be essential. It is recommended that graduate assignment lists obtained from the schools be verified with the Bureau of Naval Personnel (BUPERS).

Another problem which arose during the study concerned the fact that in many instances more than one graduate was assigned to a single supervisor. Thus, supervisors could have completed as many as three or four questionnaires (one on each graduate) or could have been interviewed three or four different times. This practice is obviously undesirable. It requires excessive amounts of supervisory time, detracts from the credibility of feedback surveys, and introduces strong possibilities of data bias from disinterested supervisors. Solutions to the problem of multiple graduates per supervisor are varied. For example, in some cases, increasing the sample size might permit supervisors to complete the survey with respect to only one specific individual or to respond in terms of an "average" graduate.

FACILITIES

PTEP reported that a variety of facilities were used for testing and interviewing AE and BT personnel. These included various types of classrooms (e.g., in hangars, at PTEP, in trailers), administrative offices, conference rooms, Chiefs' lounges, crew's quarters, log room, mess deck. Because of the small numbers of people involved in the study, it was not practicable to examine test scores or ratings obtained from the interview to determine if "place of administration" differentially affected the obtained results. It is known, however, that nonstandardized conditions of administration often do produce differential results. Both tests and SIs should be conducted under standardized conditions in private (i.e., free from interruptions), well-lighted, and otherwise comfortable areas. To the extent that they are not, one can reasonably expect differences in results. To avoid this problem in the future, attempts should be made to identify and use facilities in which certain essential features can be controlled.

PRIVACY ACT

There has been some uncertainty within military training activities regarding the applicability of the Privacy Act to surveys conducted within the military context. In the PEP study, confusion over its applicability led to some study delay. Subsequently, it was learned that the Privacy Act does not apply to surveys concerned solely with military training. Hence, this need not be a concern for such surveys.

QUESTIONNAIRE RETURN RATES

Table 16 shows the Q return rates for AE and BT graduates and supervisors. Row 1 lists the percentage of the total number of Qs returned by individuals or their units. These included Qs marked "deserted," "transferred," etc. Row 2 lists the percentage of mailout Qs returned that were usable for analysis purposes. These figures do not include usable Qs returned from respondents who also completed a JKT (N=5 for both supervisor and graduate). These respondents are excluded because they were part of a special condition added to the evaluation and hence, are not "true" mailout Q returns. The third row lists the bases over which the percentages were computed.

TABLE 16. QUESTIONNAIRE RETURN RATES FOR AE AND BT GRADUATES AND SUPERVISORS

	AE		BT	
	Graduates	Supervisors	Graduates	Supervisors
Percent Returned	41	47	50	50
Percent Usable	31	37	32	32
Base	150	150	125	125

For AEs, Q completion and return was voluntary. For BTs, Q return was virtually mandatory given COMNAVSURFLANT requirements for accountability and the instruction that Qs were to be returned to TAEG via Commanding Officers. Despite this condition, however, there was no significant difference (47 versus 50 percent) between the overall return rates of AE and BT supervisors. Fewer Qs were returned by AE graduates than by BT graduates or supervisors. The percent usable was essentially the same in all cases although AE supervisors had slightly more usable returns than the other groups of respondents. Thus, Q return rates continue to be low even when return is directed by higher authority.

Questionnaires are relatively inexpensive to construct and administer. But any large cost advantage over other methods may be lost when weighed against return rates. A number of procedures could be invoked for insuring return of a sufficient number of questionnaires such that reliable judgments about training adequacy/deficiency can be made from them. For example, completion and return could be made totally mandatory (via a number of mechanisms) with close monitoring for compliance. A second alternative which would be possible in some cases would be to increase the sample size sufficiently to compensate for anticipated low return rates.

A third way of solving the return rate dilemma is to have senior Navy enlisted personnel assigned to CNET or CNTECHTRA, dedicated to data collection, visit ships in port on a prearranged schedule. (The cognizant activity should be on distribution for ship deployment schedule information.) These personnel could assemble relatively large groups of individuals, have the Qs completed, and leave the ship with them. This tactic, in addition to insuring virtually a 100 percent "return rate," would also overcome potential biases that might arise from having questionnaires distributed and collected by Commanding Officers. Several variations on this procedure are also possible. For example, Qs could be administered either onboard or at central facilities by TRALANT or TRAPAC personnel to ship's companies in for refresher training or by ships' Training/Administrative sections.

Given the low (but not unexpected) return rate of AE Qs which were voluntary, an attempt was made to improve the return rate from BTs. This consisted of requesting Commanding Officers to oversee distribution of the Qs to the individuals from whom data was desired, assemble the completed forms, and return them to the TAEG. Also, both graduates and supervisors were instructed to return the completed forms to their Commanding Officers (within 3 weeks after receiving them) for transmittal to the TAEG. Structured interviews on the other hand were conducted by enlisted interviewers and the completed forms were taken off the ship by these interviewers.

BT DATA INCONSISTENCIES

In analyzing the BT data, it was noted that significant differences occurred in supervisor's ratings on the Q versus SI (see appendix D). BT supervisors who were interviewed rated many more tasks (N=63) as being inadequately trained than comparable supervisors who completed Qs (N=19). Also, the supervisors who were interviewed rated the proficiency of their subordinates in less than satisfactory categories for 57 BT job tasks compared to only one task rated this way by the Q group.

Since the Q and SI technique--as used in this study--are essentially the same technique (the SI is in reality an orally administered Q), the large differences were not anticipated. Note that for AEs (as shown in appendix C) differences in ratings across techniques were not as extreme.

To pinpoint the reasons for these differences, it would be necessary to conduct a further, better controlled study. Under the circumstances this was not warranted. However, several tentative hypotheses/explanations for the disparate results are possible. One of these is the essentially nonstandardized conditions under which interviews were conducted. Certain extraneous factors may have affected responses given. A second possible explanation is interviewer bias. Interviewers may have unknowingly, and unwittingly, influenced the responses given by the supervisors. There is also the possibility that the SI atmosphere--peers in a face-to-face situation--may have facilitated a frank, candid, open discussion of school training and subordinates' lack of proficiency. A fourth possible explanation, somewhat related to the above, can be found in the mechanics of Q return. As mentioned previously, BT Q respondents were instructed to return their completed Qs to their Commanding Officers for transmittal to the TAEG. This procedure may have suggested the possibility of a review by higher authority, especially to the supervisors who would probably be more sensitive to a critical review of their responses than would graduates. Consequently, they may have tended to be less critical of the school's training in the Q situation. This explanation seems especially plausible for explaining the low ratings on task proficiency obtained from the SI supervisor group. Sensitivity to the possibility of review might have precluded low ratings on subordinates by Q supervisors as this would effectively place both the subordinates and themselves on report. The SI group did not have this concern since the information given by them was taken off the ship by the interviewers.

It was previously suggested that future evaluation programs using Qs as screening devices to identify training problem areas would benefit from the use of military personnel dedicated to data collection efforts. Return rates would be greatly enhanced by this measure and it would also insure respondents that there is no possibility of review of their opinions by higher authority in their direct chain of command.

NOTAP PARTICIPATION

One of the objectives of the PEP study was to define a standardized training appraisal subsystem for the Naval Education and Training Command. An important component of this subsystem is a data processing capability for summarizing evaluation data obtained from the fleet. NOTAP "participation" was included in the PEP both to evaluate the potential utility of already existing resources for this function and to determine the general requirements for future support of training appraisal efforts.

For the AE portion of the study, NOTAP processed/prepared summaries of the SI and Q data. It was found that NOTAPs present configuration is not suitable for the data processing role in a longer-term operation. NOTAPs internal data handling procedures and data processing programs were designed for a specific purpose (i.e., summarizing task analysis data). For the AE study it was necessary for TAEG personnel--at a cost of approximately 8 man-weeks of labor--to transpose data to the booklets used by NOTAP for field task data collection. Only in this format could the data be optically scanned and automatically "scored."

Transposing data is undesirable for three reasons: (1) the extra manpower required to prepare the data for processing, (2) the undesirability of forcing data collected in one format into a format designed for a completely different purpose, and (3) the possibility of error.

Discussions were held with NOTAP concerning use of a more simple format (e.g., a standard IBM five choice machine scorable answer sheet). However, these discussions indicated that NOTAP could not readily adapt its optical scanning equipment to read these answer sheets. Also, the data processing programs used by NOTAP are not designed to accept data, process it, or output it in other than standard NOTAP formats. It should also be noted that NOTAP does not have a dedicated computer facility. NOTAP data are processed on a computer available on a restricted schedule (i.e. primarily on weekends).

Because of current NOTAP limitations and technical difficulties in coordinating data processing requirements over the telephone, TAEG elected not to use NOTAP capabilities for processing BT data. The Navy Regional Data Analysis Center (NARDAC) provided this service for the BT portion of the study. Also, the TAEG WANG programmable calculator was used to prepare selected summaries.

NOTAPs organizational mission, available data processing programs, and the lack of a readily accessible computer limits their potential role in any long-term evaluation system. If further consideration is given NOTAP in this role, resource provisions needed to implement and execute it should be obtained from/for NOTAP. Likely, these would include requirements for at least one programmer to develop data processing programs for evaluation data and for computational equipment.

For future evaluation efforts using questionnaires, it would be highly desirable for respondents to mark their choices directly on machine scorable answer sheets. These could then be run through an optical scanner tied to a computational facility. Data summaries could be processed and output on associated printing equipment. Programming services as well as the capital equipment mentioned above would be required to implement this function.

NEED FOR SINGLE POINT CONTROL

As mentioned previously, responsibilities for the PEP were apportioned among a number of participant organizations. This situation was necessitated by the realities of resource availability for the conduct of studies such as the PEP. Because of this arrangement, certain technical problems arose which are undesirable from the standpoint of "good practice" for training appraisal. For example, the requirement for data collected by one organization to be analyzed by a second organization led to difficulties in interpreting the data collected. As a case in point, the extent to which interviewers used to collect PEP data may have biased the results of the SIs is not known. Since little opportunity was provided for training these personnel in interviewing procedures and no continuous review or control over their activities was provided, it is likely that there was less than desirable standardization of their activities. For future feedback efforts, it is recommended that administrative control of personnel assigned data collection tasks be delegated to the organization responsible for analysis of that data. Similarly, it is recommended that strong attempts be made to obtain dedicated facilities, resources, and personnel for the important training appraisal function. From the PEP experience, it is clear that effective training appraisal systems and programs cannot be implemented and operated on an "out-of-hide" basis.

SECTION VI

CONCLUSIONS AND RECOMMENDATIONS

This section contains conclusions concerning the comparability of the three feedback methods evaluated. It also provides recommendations relevant to future training appraisal efforts.

CONCLUSIONS

Conclusions which can be drawn from the correlational analyses performed to assess the equivalence of data gathered by the three collection techniques are presented below.

METHOD EQUIVALENCE. With respect to the equivalence of the methods, the correlational analyses support the following conclusions:

- Equivalent frequency of task performance data can be obtained by either the Q or the SI method.
- The Q and SI yield equivalent ratings of training adequacy.
- Assessment of proficiency can be accomplished using either the Q or SI method.
- Because JKT scores do not correlate with ratings of training adequacy or proficiency, it must be concluded that the SI and JKT methods do not produce equivalent information about training effects.
- The lack of correlation between JKT scores and SI ratings prevents any inferences concerning the equivalency of the Q and JKT methods.

In summary, the Q and SI methods used in this study are fully interchangeable for data collection. The JKT and Q/SI methods produce different information.

SOURCE EQUIVALENCE. The correlational analyses performed to assess the equivalence of data sources support the general conclusion that comparable frequency of task performance and training adequacy information can be obtained from either graduates or supervisors.

INDEPENDENCE OF SCALES. All combinations of ratings of training adequacy with supervisor ratings of proficiency were significantly correlated. This suggests that one of these two scales can be eliminated in future training appraisal surveys. Information presented in this report supports the conclusion that the proficiency scale not be used when the purpose of assessment is to evaluate the quality of training.

Firm conclusions regarding relationships between the frequency scale and/or the training adequacy and proficiency scales could not be established because of inconsistencies between the AE and BT data. The decision to use a

"frequency of task performance" scale in future training appraisal surveys, however, should be tempered by the potential uses and value of such information. In many instances, frequency of task performance information may not be as meaningful for curriculum review purposes as information bearing more directly on the importance of including training for particular tasks in a school's curriculum. Also, frequency of task performance is often set by established maintenance procedures. Consequently, frequency information, if needed, can be obtained more economically and, perhaps, more reliably from published sources.

RESOURCE COMPARISONS. The results of the analyses of PEP resource data demonstrate that, overall, the Q method is least expensive for collecting data. The SI method is more costly than the Q method because of the one-on-one assessment situation. Primarily, because of high developmental costs, the JKT is the most expensive technique. These costs are directly proportional to the number of tests administered with costs of a given JKT decreasing as the number of tests increase. For the same number of individuals "tested," total SI costs would increase because of the two-person situation. Mailout Q costs will vary as a function of postage and reproduction costs.

The final choice of a data collection method, however, should not be based solely on resource considerations. As discussed previously, the instrument used must consider fully the information required from an appraisal effort.

RECOMMENDATIONS

Recommendations for the conduct of future training appraisal efforts are presented below. These are based on work accomplished and experience acquired during the PEP. The following recommendations with accompanying rationales are offered:

1. It was demonstrated in this study that graduates and supervisors provide equivalent data concerning frequency of task performance and adequacy of "A" school training. Under the usual conditions of limited resources and restricted availability of personnel, it is recommended that choice of data source be determined by convenience of access, costs, and/or purpose of assessment. In many cases, supervisors would be preferred because of their greater experience which can be expected to result in more reliable data.

2. The results of this study demonstrate that the Q and SI methods used here produce equivalent data while the JKT provides different information than these two methods. However, discussions provided in the text support conclusions that each of these methods has differential utility for satisfying particular information needs. Based on those discussions, the following recommendations are made:

a. Low cost questionnaires of the type used in this study are recommended for use as screening devices to determine the presence (or absence) of deficiencies in the quality of school training.

b. Structured interview procedures are recommended when more detailed information is required concerning the specific nature of training deficiencies. Note that the SI procedure used in this study was not designed to provide such information.

c. Job knowledge testing should be restricted to the purposes for which it is intended; i.e., to assess trainee possession of factual knowledge of job/equipment operations and repair. In its current state of development and application, it seems inappropriate for field use as a technique for identifying deficiencies in school training.

3. Mechanisms for accessing the fleet to obtain data in a timely manner (i.e., without lengthy delays) should be established. A discussion of the relative merits of possible approaches is not warranted here since separate study should be devoted to evaluating the various alternatives. Some suggestions in this area, however, are:

- written agreements with the cognizant fleet user(s) of NAVEDTRACOM graduates,
- CNO instructions or directives requiring command cooperation in feedback efforts conducted by the NAVEDTRACOM either on an as-needed basis or during "set aside" periods,
- feedback gathering by fleet units with results supplied to the TRACOM for action,
- implementation of a training appraisal system with periodic appraisal of all schools. This would operate on a routine basis such that advance planning/coordination with fleet units would not be required on a case-by-case basis.

4. If individual students are to be the focus of future appraisal efforts, then care should be taken to insure that the graduates' units of assignment are correctly identified. Information obtained from the schools should be verified with BUPERS prior to composing samples or attempting to contact specific individuals. Also, survey procedures should be such that any given supervisor should be requested to complete only one feedback survey.

5. At least for those survey efforts which involve the use of tests or interviews for gathering feedback information, some attempt should be made to obtain suitable facilities in which data gathering conditions can be standardized.

6. Steps should be taken to improve return rates from questionnaires. Several means for doing this were suggested in the text. The most promising seems to be one in which Naval enlisted personnel administer Qs to groups of "anonymous" individuals and leave the ship(s) with the completed Qs. In addition to insuring a high number of returns, this procedure would also guard against potential response bias that could be introduced by permitting command review of the questionnaires.

7. More efficient data handling and processing facilities/capabilities will be required for compiling and summarizing results of future evaluation surveys. It is recommended that questionnaires either be preprinted on machine scorable answer sheets or such sheets be used in conjunction with questionnaires. These sheets could then be optically scanned on equipment tied to a computer. Data summaries could then be directly output on printing equipment.

REFERENCES

Central Test Site for the Personnel and Training Evaluation Program, PTEP Evaluation 7-78, Boiler Technician (1200 PSI) Job Knowledge Test Report. December 1978. Central Test Site For the Personnel and Training Evaluation Program. Naval Guided Missile School. Dam Neck, VA 23461

Central Test Site for the Personnel and Training Evaluation Program, PTEP Evaluation 8-77, Aviation Electrician's Job Knowledge Test Report. October 1977. Central Test Site for the Personnel and Training Evaluation Program. Naval Guided Missile School, Dam Neck, VA 23461

Denton, Carol F. and Hall, E. R. Boiler Technician "A" School (1200 PSI) Training Assessment Data, TAEG Technical Memorandum 78-1. July 1978. Training Analysis and Evaluation Group, Orlando, FL 32813

Dyer, F. N., Ryan, L. E. and Mew, Dorothy V. A Method for Obtaining Post Formal Training Feedback: Development and Evaluation. TAEG Report No. 19. May 1975. Training Analysis and Evaluation Group, Orlando, FL 32813

Guilford, J. P. Psychometric Methods, New York: McGraw-Hill, Inc. 1954

Hall, E. R. and Denton, Carol F. Aviation Electrician's Mate "A" School Training Assessment Data, TAEG Technical Memorandum 77-9. November 1977. Training Analysis and Evaluation Group, Orlando, FL 32813

Hall, E. R., Rankin, W. C. and Aagard, J. A. Training Effectiveness Assessment: Volume II, Problems, Concepts and Evaluation Alternatives. TAEG Report No. 39. December 1976. Training Analysis and Evaluation Group, Orlando, FL 32813

APPENDIX A

SPECIFIC JOB TASKS RATED BY AEs

The 124 job tasks rated by the AEs are provided in this appendix. The alphanumeric code in parentheses is a cross reference to the NOTAP AE job Task Inventory.

SPECIFIC JOB TASKS RATED BY AEs

1. (K14) CONNECT EXTERNAL POWER TO AIRCRAFT
2. (K31) INSTALL AND/OR REMOVE AIRCRAFT PROTECTIVE COVERS (e.g., PITOT-STATIC TUBE COVER)
3. (L1) SAFETY WIRE EQUIPMENT, GEAR, AND SWITCHES
4. (K8) TIGHTEN LOOSE SCREWS, CAM-LOCK FASTENERS, ETC.
5. (L2) ADJUST MECHANICAL LINKAGES (e.g., MICROSWITCH RIGGING)
6. (L9) FABRICATE AND REPAIR INTERCONNECTING ELECTRICAL CABLES
7. (L10) REPLACE PINS IN ELECTRICAL CONNECTORS (e.g., CANNON PLUGS)
8. (L11) REPLACE "O" RINGS, GASKETS, AND SEALS
9. (L12) REPLACE KNOBS, LIGHTS, FUSES, AND CIRCUIT BREAKERS
10. (K13) REPLACE COMMON HARDWARE (e.g., NUTS, BOLTS, AND SCREWS)
11. (L13) POT RELAYS, PLUGS, POWER SUPPLIES, SWITCHES, ETC.
12. (L14) REPLACE ELECTRICAL CONNECTORS (e.g., CANNON PLUGS)
13. (L18) BUILD-UP, FABRICATE, AND REPAIR COAXIAL CABLES
14. (L19) SPLICE ELECTRICAL WIRES
15. (L20) SPOT-TIE WIRING
16. (L29) REPLACE BONDING WIRE
17. (L36) APPLY RANGE MARKS TO INSTRUMENTS
18. (L38) CLEAN RELAY AND SWITCH CONTACTS
19. (L51) CLEAR PITOT-STATIC SYSTEM OF WATER AND OTHER FOREIGN MATTER
20. (M4) TEST CIRCUITS FOR PROPER PHASE SEQUENCE
21. (P1) USE BLOCK DIAGRAMS
22. (P3) USE WIRING DIAGRAMS
23. (P2) USE SCHEMATIC DIAGRAMS
24. (P6) USE PUBLICATIONS FOR TROUBLESHOOTING ASSISTANCE AND PERFORMANCE DATA

SPECIFIC JOB TASKS RATED BY AEs (continued)

25. (F1) USE PUBLICATIONS TO OBTAIN SUPPLY DATA AND ORDER PARTS AND SUPPLIES
26. (M5) TEST WIRING AND COAXIAL CABLE USING MEGGER
27. (P10) VISUALLY INSPECT EQUIPMENT FOR DEFECTS
28. (P11) TAKE VOLTAGE READINGS
29. (P14) MEASURE RESISTANCE
30. (P13) PERFORM CONTINUITY CHECKS
31. (N58) CLEAN AIRCRAFT BATTERIES
32. (P15) MEASURE CAPACITANCE OTHER THAN FUEL QUANTITY SYSTEM
33. (P16) MEASURE FREQUENCY
34. (P17) TAKE ELECTRICAL CURRENT READINGS
35. (N1) PERFORM AIRCRAFT PHASE INSPECTION
36. (N2) PERFORM AIRCRAFT ACCEPTANCE AND TRANSFER INSPECTIONS
37. (N3) PERFORM DAILY INSPECTIONS
38. (N4) PERFORM PREFLIGHT INSPECTIONS
39. (N5) PERFORM POSTFLIGHT INSPECTIONS
40. (N17) PERFORM AIRCRAFT CONDITIONAL INSPECTIONS
41. (N19) INSPECT EQUIPMENT FOR SECURITY AND PROPER INSTALLATION
42. (G4) CONDUCT GENERAL SAFETY INSPECTION OF EQUIPMENT PRIOR TO USE (e.g., TEST EQUIPMENT, TOOLS)
43. (G5) CONDUCT COCKPIT SAFETY INSPECTION PRIOR TO MAINTENANCE
44. (G6) COMPLY WITH SAFETY MESSAGES, PROCEDURES, AND DIRECTIVES
45. (G3) INSPECT AIRCRAFT ELECTRICAL COMPONENTS FOR CORROSION IN ACCORDANCE WITH APPLICABLE MAINTENANCE REQUIREMENTS
46. (M10) TEST, CHECK, AND TROUBLESHOOT ELECTRICALLY ACTUATED HYDRAULIC VALVES
47. (M11) TEST, CHECK, AND TROUBLESHOOT ACCELEROMETERS
48. (M16) TEST, CHECK, AND TROUBLESHOOT FLIGHT CONTROL SYSTEM

SPECIFIC JOB TASKS RATED BY AEs (continued)

49. (M14) TEST, CHECK, AND TROUBLESHOOT FLIGHT CONTROL TRIM SYSTEM
50. (M18) TEST, CHECK, AND TROUBLESHOOT WING ANTI-ICE SYSTEM
51. (M19) TEST, CHECK, AND TROUBLESHOOT WINDSHIELD DE-ICE SYSTEM
52. (M25) TEST, CHECK, AND TROUBLESHOOT TURN AND BANK INDICATORS
53. (M26) TEST, CHECK, AND TROUBLESHOOT TRUE AIRSPEED INDICATING SYSTEM
54. (M27) TEST, CHECK, AND TROUBLESHOOT ATTITUDE INDICATING SYSTEM
55. (M28) TEST, CHECK, AND TROUBLESHOOT AIRSPEED INDICATORS
56. (M29) TEST, CHECK, AND TROUBLESHOOT AIRCRAFT QUANTITY INDICATING SYSTEMS
57. (M31) TEST BATTERIES FOR SPECIFIC GRAVITY (USING A HYDROMETER)
58. (M33) TEST, CHECK, AND TROUBLESHOOT EMERGENCY GENERATORS (RAT, EPP)
59. (M34) TEST, CHECK, AND TROUBLESHOOT GENERATORS
60. (M35) TEST, CHECK, AND TROUBLESHOOT TRANSFORMER RECTIFIER SYSTEM
61. (M38) TEST, CHECK, AND TROUBLESHOOT VOLTAGE REGULATORS
62. (M45) TEST, CHECK, AND TROUBLESHOOT ENGINE START SYSTEM (ELECTRICAL)
63. (M42) TEST, CHECK, AND TROUBLESHOOT ENGINE IGNITION (IGNITER) SYSTEM
64. (Q9) TEST, CHECK, AND TROUBLESHOOT ELECTRO-HYDRAULIC SYSTEMS
65. (M46) TEST, CHECK, AND TROUBLESHOOT AIRCRAFT LIGHTING SYSTEM (INTERIOR)
66. (M47) TEST, CHECK, AND TROUBLESHOOT AIRCRAFT LIGHTING SYSTEM (EXTERIOR)
67. (M48) TEST, CHECK, AND TROUBLESHOOT ENGINE FUEL PRESSURE INDICATING SYSTEM
68. (M50) TEST, CHECK, AND TROUBLESHOOT ENGINE FUEL INDICATING SYSTEM
69. (M51) TEST, CHECK, AND TROUBLESHOOT OIL PRESSURE INDICATING SYSTEM
70. (M52) TEST, CHECK, AND TROUBLESHOOT OIL TEMPERATURE INDICATING SYSTEM
71. (M58) TEST, CHECK, AND TROUBLESHOOT HYDRAULIC PRESSURE INDICATING SYSTEM
72. (M61) TEST, CHECK, AND TROUBLESHOOT AUTOPILOT SYSTEM
73. (Q26) TEST, CHECK, AND TROUBLESHOOT COMPONENTS OF AIRCRAFT ELECTRICAL POWER SYSTEMS

SPECIFIC JOB TASKS RATED BY AEs (continued)

74. (Q13) TEST, CHECK, AND TROUBLESHOOT NOSE WHEEL STEERING SYSTEM

75. (Q10) TEST, CHECK, AND TROUBLESHOOT POSITION INDICATING SYSTEM (e.g., GEAR, FLAPS)

76. (R38) REMOVE AND REPLACE COMPONENTS OF AIRCRAFT WARNING INDICATOR SYSTEM

77. (R26) REMOVE AND REPLACE COMPONENTS OF AIRCRAFT QUANTITY INDICATOR SYSTEMS

78. (R27) REMOVE AND REPLACE COMPONENTS OF AIRCRAFT ELECTRICAL POWER SYSTEM

79. (R2) REPLACE AIRCRAFT ELECTRICAL WIRING

80. (R12) REMOVE AND/OR REPLACE PRINTED CIRCUIT CARDS, BOARDS, AND MODULES

81. (R3) REPLACE ELECTRIC MOTORS

82. (R18) REMOVE AND REPLACE COMPONENTS OF AIRCRAFT ANTI-ICE/DE-ICE SYSTEM

83. (R20) REMOVE AND REPLACE COMPONENTS OF AIRCRAFT FLIGHT INSTRUMENT SYSTEM

84. (R16) CHANGE PILOT'S CONTROL STICK

85. (R22) CHANGE AIRCRAFT ACCELEROMETERS

86. (R23) CHANGE AIRCRAFT CLOCKS

87. (R24) CHANGE STANDBY COMPASS

88. (R25) CHANGE AIRCRAFT PITOT-STATIC LINES

89. (R19) CHANGE PITOT-STATIC TUBES

90. (L5) CHANGE ELECTRICAL COMPONENTS OF NOSE WHEEL STEERING

91. (R28) CHANGE AIRCRAFT BATTERIES

92. (R30) CHANGE AIRCRAFT EXTERNAL POWER RECEPTACLES

93. (R35) CHANGE AIRCRAFT GENERATORS

94. (R39) CHANGE AIRCRAFT FIRE WARNING ELEMENTS/THERMOCOUPLES

95. (R42) REMOVE AND REPLACE COMPONENTS OF AIRCRAFT LIGHTING SYSTEMS

SPECIFIC JOB TASKS RATED BY AEs (continued)

96. (R43) REMOVE AND REPLACE COMPONENTS OF AIRCRAFT ENGINE PERFORMANCE INDICATOR SYSTEM

97. (R47) REMOVE AND REPLACE COMPONENTS OF PRESSURE INDICATING SYSTEM

98. (R49) REMOVE AND REPLACE COMPONENTS OF AUTOFLIGHT CONTROL/AUTO PILOT SYSTEM

99. (S4) ADJUST SYNCHROS AND RESOLVERS

100. (S5) ADJUST SERVO AMPLIFIERS

101. (S13) ADJUST MICROSWITCHES

102. (W24) ADJUST AUTOMATIC FLIGHT CONTROL/AUTO PILOT SYSTEM

103. (W9) ADJUST VOLTAGE REGULATOR

104. (W14) ADJUST PRESSURE INDICATING SYSTEMS

105. (S15) ADJUST THERMOSTATS AND THERMOSWITCHES

106. (S16) ADJUST VARIABLE RESISTORS

107. (T9) RIG FLIGHT CONTROL TRIM SYSTEM

108. (W2) ADJUST ANGLE-OF-ATTACK INDICATOR

109. (W3) ADJUST RATE-OF-CLIMB INDICATOR

110. (W4) ADJUST VERTICAL GYRO INDICATOR AND HORIZONTAL INDICATOR

111. (Z25) ADJUST NOSE WHEEL STEERING

112. (W5) SWING AND COMPENSATE AIRCRAFT COMPASS SYSTEM

113. (W6) SET BAROMETRIC ALTIMETER SCALE TO FIELD ELEVATION

114. (W7) CALIBRATE AIRCRAFT FUEL QUANTITY INDICATING SYSTEM

115. (W17) CALIBRATE ENGINE TEMPERATURE INDICATING SYSTEM

116. (G7) SIGN FOR CLASSIFIED MATERIAL

117. (G8) FILE AND STOW CLASSIFIED MATERIAL

118. (G9) DESTROY CLASSIFIED MATERIAL

119. (H1) FILL OUT MAINTENANCE SUPPORT ACTION FORMS (VIDS/MAF)

120. (H7) SCREEN VIDS/MAF CARDS FOR COMPLETENESS AND CORRECTIONS

SPECIFIC JOB TASKS RATED BY AEs (continued)

121. (H2) FILL OUT SUPPORT ACTION FORMS

122. (H3) FILL OUT TECHNICAL DIRECTIVE COMPLIANCE FORMS

123. (H4) PREPARE ROUGH DRAFTS OF UNSATISFACTORY REPORTS

124. (H5) MAINTAIN VISUAL INFORMATION DISPLAY (VIDS) SYSTEM BOARDS

APPENDIX B

SPECIFIC JOB TASKS RATED BY BTs

A complete listing of the 124 job tasks which were rated by BTs is provided in this appendix. The identifiers used in parentheses are cross references to the NOTAP BT Job Task Inventory printout or to watch station tasks provided by the BT "A" School or to Personnel Advancement Requirements.

The asterisk identifies tasks which were not taught in BT "A" School at the time the survey instruments were devised.

TAEG Report No. 64

SPECIFIC JOB TASKS RATED BY BTs

1. (G15) USE ENGINEERING OPERATIONS SEQUENCING SYSTEM (EOSS) MANUAL
2. (OC7) PERFORM BOILER PRELIGHT-OFF INSPECTIONS
3. (P29) LINE UP/SECURE MAIN FEED RECIRCULATION SYSTEM
4. (S1) LINE UP FUEL OIL HEATER DRAINS IN ACCORDANCE WITH EOSS
5. (P36) LINE UP/SHIFT TO-FROM LP/HP DRAIN SYSTEMS
6. (P57) LINE UP MACHINERY COOLING WATER REDUCING STATION
7. (S2) LINE UP FRESH WATER DRAIN COLLECTING TANK SYSTEM IN ACCORDANCE WITH EOSS
8. (S3) LINE UP FRESH WATER DRAIN COLLECTION SYSTEM TO DEAERATING FEED TANK IN ACCORDANCE WITH EOSS
9. (S4) LINE UP EXCESS FEED SYSTEM IN ACCORDANCE WITH EOSS
10. (S5) LINE UP/MAKE FEED SYSTEM IN ACCORDANCE WITH EOSS
11. (S6) LINE UP FEED WATER FILLING AND TRANSFER SYSTEM IN ACCORDANCE WITH EOSS
12. (P30) LINE UP STEAM SYSTEMS (MAIN AND AUXILIARY)
13. (P30) LINE UP SUPERHEATER PROTECTION SYSTEM IN ACCORDANCE WITH EOSS
14. (P30) LINE UP AUXILIARY EXHAUST STEAM SYSTEM IN ACCORDANCE WITH EOSS
15. (S8) LINE UP AUXILIARY GLAND EXHAUSTER SYSTEM IN ACCORDANCE WITH EOSS
16. (P34) LINE UP/SECURE STEAM REDUCING SYSTEMS
17. (P34) LINE UP/SECURE STEAM AUGMENTING SYSTEM
18. (P54) OPEN/CLOSE ECONOMIZER VENT AND AIR COCKS
19. (S7) LINE UP LOW PRESSURE AIR SYSTEMS IN ACCORDANCE WITH EOSS
20. (P36) LINE UP/SECURE MAIN FEED SYSTEM
21. (P24) MAKE UP LIGHTING OFF BURNER
22. (UL6) START/STOP MOTOR DRIVEN LIGHTING OFF FORCED DRAFT BLOWER IN ACCORDANCE WITH EOSS
23. (P56) WARM UP/PLACE IN OPERATION/SECURE DEAERATING FEED TANK

SPECIFIC JOB TASKS RATED BY BTs (continued)

24.	(UL2)	BLOW DOWN BOILER GAUGE GLASS IN ACCORDANCE WITH EOSS
25.*	(S1)	LINE UP MOTOR DRIVEN FUEL OIL SERVICE PUMP IN ACCORDANCE WITH EOSS
26.	(P11)	LIGHT BOILER FIRES MANUALLY
27.	(B1)	LIGHT FIRES WITH ASSISTANCE WHEN ATOMIZING STEAM IS AVAILABLE IN ACCORDANCE WITH EOSS
28.*	(P8)	LIGHT OFF UNDER NO STEAM CONDITIONS
29.	(P20)	REGULATE FORCED DRAFT BLOWER SPEED FOR PROPER COMBUSTION
30.	(UL6)	LINE UP/START/SECURE TURBINE DRIVEN FORCED DRAFT BLOWER
31.	(P5)	CONTROL PROPER BOILER WATER LEVEL (LIGHTING OFF, STEAMING, SECURING)
32.	(UL2)	SECURE STEAM BLANKET SYSTEM ON MAIN BOILER IN ACCORDANCE WITH EOSS
33.*	(G13)	USE LIGHTING OFF/SECURING SHEET
34.	(P35)	LINE UP/SECURE IN-LINE DESUPERHEATER
35.	(P26)	MAKE UP ATOMIZERS
36.	(P58)	PURGE BOILER FIRE BOX
37.*	(P1)	MANUALLY CONTROL BOILER FIRING RATE
38.*	(P2)	CONTROL FIRING RATE OF SINGLE BOILER IN REMOTE MANUAL
39.*	(P3)	CONTROL FIRING RATE OF SINGLE BOILER IN LOCAL MANUAL
40.	(P4)	MONITOR FIRING RATE OF SINGLE BOILER IN AUTOMATIC
41.	(P17)	LINE UP SYSTEM FOR BOILER SURFACE/BOTTOM BLOW DOWN
42.	(P18)	OPERATE SOOT BLOWERS (AUTOMATIC)
43.	(P19)	OPERATE SOOT BLOWERS (MANUAL)
44.	(OC9)	OPERATE PERISCOPE SMOKE INDICATOR (IDENTIFY STACK CONDITIONS)
45.	(P27)	ADJUST BURNER TIP SETTINGS
46.	(P6)	OPEN/CLOSE MAIN AND AUXILIARY STEAM STOPS
47.	(P7)	PLACE MACHINERY ON REGULATING GOVERNORS

*Tasks not taught in BT "A" School

SPECIFIC JOB TASKS RATED BY BTs (continued)

48.	(H12)	MANUALLY LIFT RELIEF VALVES
49.*	(Q3)	DRAW FEED WATER SAMPLES
50.*	(Q1)	DRAW BOILER WATER SAMPLES
51.*	(P51)	SOUND FEED WATER, FRESH WATER, AND FUEL OIL TANKS AND RECORD READINGS
52.	(K27)	MONITOR AND RECORD RPM READINGS
53.	(UL1)	MONITOR PRESSURE, TEMPERATURE, SPEED INDICATING DEVICES OF OPERATING MACHINERY
54.	(G12)	MAKE ENTRIES ON BOILER ROOM OPERATING RECORD
55.	(Q12)	DRAW SAMPLES OF LUBE OIL AND CHECK FOR PRESENCE OF WATER AND/OR SEDIMENT
56.	(P50)	SHIFT DUPLEX STRAINERS (LUBE OIL/FUEL OIL)
57.	(J7)	CLEAN STRAINERS (e.g., SALT WATER/LUBE OIL/FUEL OIL)
58.	(P25)	LINE UP BOILER CHEMICAL INJECTION SYSTEM
59.	(Q10)	INJECT CHEMICALS INTO BOILER WATER TO KEEP BOILER CHEMISTRY WITHIN PRESCRIBED LIMITS
60.*	(P42)	LINE UP MAIN DRAINAGE SYSTEM TO PUMP BILGES
61.*	(P43)	LINE UP/SECURE FUEL OIL TANK STRIPPING PUMP
62.	(D18)	PARTICIPATE AS WATCHSTANDER IN BASIC ENGINEERING CASUALTY CONTROL EXERCISES (BECCE) DRILLS
63.*	(D19)	PARTICIPATE AS WATCHSTANDER IN BASIC DAMAGE CONTROL EXERCISES (BDCE)
64.	(UL5)	SHIFT AUTOMATIC FEED WATER CONTROL VALVE FROM MANUAL TO AUTOMATIC CONTROL IN ACCORDANCE WITH EOSS
65.	(UL6)	LINE UP FEED WATER FILLING AND TRANSFER SYSTEM IN ACCORDANCE WITH THE EOSS
66.	(P13)	PERFORM FIREROOM EMERGENCY WRAP-UP (MANUALLY)
67.	(P14)	SECURE BOILER FIRES MANUALLY
68.	(UL3)	PLACE STEAM BLANKET ON BOILER IN ACCORDANCE WITH EOSS

*Tasks not taught in BT "A" School

SPECIFIC JOB TASKS RATED BY BTs (continued)

69. (UL4) PLACE NITROGEN BLANKET ON BOILER IN ACCORDANCE WITH EOSS

70. (UL4) SECURE NITROGEN BLANKET ON BOILER IN ACCORDANCE WITH EOSS

71. (P16) LAY UP BOILER (DRY)

72. (OC10) USE MAINTENANCE REQUIREMENT CARDS (MRCs)

73. (OC11) PREPARE MAINTENANCE DATA COLLECTION (MDC) FORMS

74. (G29) USE TECHNICAL AND MAINTENANCE MANUALS

75.* (J23) SAFETY TAG VALVES/SWITCHES

76.* (H8) INSPECT CONDITION OF SOOT BLOWERS

77.* (H9) CHECK SOOT BLOWER SWEEPING ARCS

78.* (H10) CHECK AND ADJUST SOOT BLOWER OPERATING PRESSURE

79.* (H7) REPLACE BURNER GASKETS

80. (N1) CLEAN BURNER TIPS

81. (N3) CLEAN BURNER BARRELS

82. (OC6) PERFORM PREVENTIVE MAINTENANCE ON PUMPS

83. (K7) REPLACE PUMP GASKETS

84. (K8) REMOVE/REPLACE PUMP SEALS/PACKING

85.* (K13) ASSIST WITH ALIGNMENT OF PUMP DRIVING UNIT

86.* (K15) REMOVE/REPLACE FLEXIBLE COUPLINGS

87.* (K14) CLEAN/LUBRICATE FLEXIBLE COUPLINGS

88. (K16) INSPECT CONDITION OF FLEXIBLE COUPLINGS

89.* (K26) TAKE THRUST READINGS

90 * (L12) ASSIST IN SETTING/TESTING OF SPEED LIMITING GOVERNORS

91.* (L14) TEST OVERSPEED TRIPS

92.* (L15) CHECK/SET LOW PRESSURE TRIP DEVICE ON MAIN FEED PUMP

93.* (J24) BACK FLUSH LUBE OIL COOLER

*Tasks not taught in BT "A" School

SPECIFIC JOB TASKS RATED BY BTs (continued)

94.* (K19) CLEAN LUBE OIL SUMPS OF AUXILIARY MACHINERY

95.* (K20) CHANGE LUBE OIL AUXILIARY MACHINERY

96. (P29) LINE UP/SECURE FUEL OIL SERVICE SYSTEM

97.* (OC8) PERFORM HYDROSTATIC TEST ON LUBE OIL COOLER

98.* (H20) PERFORM HYDROSTATIC TEST ON DESUPERHEATER

99. (H23) PRESERVE/GREASE BOILER SLIDING FEET

100. (H26) REMOVE/REPLACE BOILER ACCESS DOORS

101. (H42) ASSIST IN HYDROSTATIC TEST OF BOILERS

102. (H50) REMOVE/CLEAN/REPLACE HAND HOLD SEATS/GASKETS

103. (J1) SOUND/CHECK FOUNDATION BOLTS

104.* (H29) APPLY METAL CONDITIONING COMPOUND TO FIRESIDE

105. (H38) REMOVE/REPLACE STEAM DRUM INTERNAL FITTINGS

106. (N4) CLEAN/PRESERVE BOILER AIR CASING

107. (N5) CLEAN FORCED DRAFT BLOWER DUCTS

108. (N6) CLEAN FIRESIDES

109. (N7) CLEAN WATERSIDES

110. (N8) CLEAN OUT EXPANSION JOINTS

111. (N10) CLEAN LIQUID LEVEL AND/OR FLOW SIGHT GLASSES

112. (J3) REMOVE/REPLACE FILTERS

113. (J6) MANUFACTURE GASKETS FOR VALVES/BLANK FLANGES, ETC.

114. (J15) INSTALL BLANK FLANGES ON PIPING

115. (J10) REPAIR/REPLACE LAGGING

116. (M18) HAND GRIND/SPOT-IN VALVES

117. (R1) REMOVE/REPLACE BOILER GAUGE GLASSES

118. (OC1) OPERATE DEAD WEIGHT TESTER

*Tasks not taught by BT "A" School

SPECIFIC JOB TASKS RATED BY BTs (continued)

119. (K34) CALIBRATE PRESSURE GAUGES USING DEAD WEIGHT TESTER OR GAUGE COMPARATOR

120. (K35) REMOVE/REPLACE PRESSURE/TEMPERATURE GAUGES

121. (OC2) MEASURE SHAFT RPM's USING STROBOSCOPE

122. (M17) TEST COMBINATION EXHAUST RELIEF VALVES

123. (OC3) OPERATE DISTILLING PLANT

124. (C17) FILL OUT WORK REQUEST/WORK ORDERS

APPENDIX C-1

TASKS RATED WITHIN SPECIFIED SUBCATEGORIES OF FREQUENCY SCALE BY AEs

Response data for the frequency of task scale are given in this appendix without regard to aircraft type or maintenance level. The data presented in table C-1 are summaries of the number of tasks rated within the five subcategories by instrument (questionnaire (Q) and structured interview (SI)) and by respondent (graduates (G) and supervisors (S)). The number of tasks on which graduates and supervisors agreed in their subcategory assignments is also shown. The final column lists the number of tasks on which all respondents agreed regardless of instrument used. The numbers in parentheses identify the mean values for those tasks rated in these subcategories.

TABLE C-1. NUMBER OF TASKS RATED WITHIN SPECIFIED FREQUENCY SUBCATEGORIES BY GRADUATES (G) AND SUPERVISORS (S)

Subcategories	Q			SI			All Agree
	G	S	Agree	G	S	Agree	
1.00 - 1.49 Never Performed	3 (1.29)	13 (1.31)	3	14 (1.28)	8 (1.22)	5	3
1.50 - 2.49 Seldom Performed	61 (1.98)	56 (1.97)	46	48 (1.97)	56 (2.04)	41	32
2.50 - 3.49 Performed Monthly	33 (2.93)	29 (2.83)	24	34 (2.87)	33 (2.98)	27	18
3.50 - 4.49 Performed Weekly	21 (4.18)	16 (4.09)	17	25 (4.08)	14 (4.01)	12	11
4.50 - 5.00 Performed Daily	6 (4.59)	10 (4.64)	6	3 (4.80)	13 (4.67)	3	1

Also shown in this appendix are lists of the identity numbers of the tasks rated by AEs, grouped in the same mean value subcategories. The tasks are grouped by respondent, instrument, and commonality of ratings in terms of the instrument. The meanings ascribed to the mean values are provided as category subheadings.

1.00 - 1.49 NEVER PERFORMED
Graduate - Q TASKS: 105, 117, 118
Supervisor - Q TASKS: 50, 57, 91, 92, 103, 104, 105, 107, 112, 115, 116, 117, 118
Common - Q TASKS: 105, 117, 118
Graduate - SI TASKS: 32, 50, 57, 63, 88, 92, 103, 104, 105, 109, 110, 112, 117, 118
Supervisor - SI TASKS: 38, 39, 57, 105, 110, 116, 117, 118
Common - SI TASKS: 57, 105, 110, 117, 118
Common for Respondents on both Q and SI TASKS: 105, 117, 118

1.50 - 2.49 SELDOM PERFORMED
Graduate - Q TASKS: 2, 8, 17, 20, 31, 32, 33, 36, 38, 39, 47, 50, 51, 52, 57, 58, 59, 60, 61, 62, 63, 67, 68, 69, 70, 71, 74, 80, 81, 82, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 96, 97, 99, 100, 101, 103, 104, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 122, 123
Supervisor - Q TASKS: 16, 17, 20, 26, 31, 32, 33, 34, 36, 38, 39, 40, 47, 51, 52, 53, 58, 59, 60, 61, 62, 63, 67, 68, 69, 70, 71, 74, 78, 79, 80, 81, 82, 84, 85, 86, 87, 88, 89, 90, 93, 94, 96, 97, 99, 100, 102, 106, 108, 109, 110, 111, 113, 114, 122, 123
Common - Q TASKS: 17, 20, 31, 32, 33, 36, 38, 39, 47, 51, 52, 58, 59, 60, 61, 62, 63, 67, 68, 69, 70, 71, 74, 80, 81, 82, 84, 85, 86, 87, 88, 89, 90, 93, 94, 96, 97, 99, 100, 106, 108, 109, 110, 111, 113, 114
Graduate - SI TASKS: 8, 13, 16, 20, 26, 31, 33, 36, 38, 39, 40, 47, 51, 52, 53, 58, 59, 60, 61, 62, 67, 68, 69, 70, 71, 74, 82, 84, 85, 86, 87, 89, 90, 93, 94, 99, 100, 101, 106, 107, 108, 111, 113, 114, 115, 116, 122, 123
Supervisor - SI TASKS: 13, 16, 17, 20, 26, 31, 32, 33, 34, 36, 37, 47, 50, 51, 52, 53, 58, 59, 60, 61, 62, 63, 67, 68, 69, 70, 71, 74, 81, 82, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 99, 100, 103, 104, 106, 107, 108, 109, 111, 112, 113, 114, 115, 122, 123
Common - SI TASKS: 13, 16, 20, 26, 31, 33, 36, 47, 51, 52, 53, 58, 60, 61, 67, 68, 69, 70, 71, 74, 82, 84, 85, 86, 87, 89, 90, 93, 94, 99, 100, 106, 107, 108, 111, 113, 114, 115, 116, 122, 123
Common for Respondents on both Q and SI TASKS: 20, 31, 33, 36, 47, 51, 52, 58, 60, 61, 67, 68, 69, 70, 71, 74, 82, 84, 85, 86, 87, 89, 90, 93, 94, 99, 100, 106, 108, 111, 113, 114

2.50 - 3.49 PERFORMED MONTHLY
Graduate - Q TASKS: 5, 6, 7, 11, 12, 13, 16, 18, 19, 26, 34, 35, 37, 40, 46, 48, 49, 53, 54, 55, 56, 64, 72, 73, 75, 76, 77, 78, 79, 83, 95, 98, 102
Supervisor - Q TASKS: 2, 5, 6, 7, 8, 11, 12, 13, 18, 19, 35, 37, 46, 48, 54, 55, 56, 64, 72, 73, 75, 76, 77, 83, 95, 98, 101, 124
Common - Q TASKS: 5, 6, 7, 11, 12, 13, 18, 19, 35, 37, 46, 48, 49, 55, 56, 64, 72, 73, 75, 76, 77, 83, 95, 98
Graduate - SI TASKS: 2, 5, 6, 7, 11, 12, 17, 18, 19, 34, 35, 37, 46, 49, 54, 55, 56, 64, 72, 73, 75, 76, 77, 78, 79, 80, 81, 91, 96, 97, 98, 102, 121, 124
Supervisor - SI TASKS: 2, 5, 6, 7, 8, 11, 12, 18, 19, 35, 40, 44, 46, 49, 54, 55, 56, 64, 73, 75, 76, 77, 78, 79, 80, 83, 95, 96, 97, 98, 101, 102, 124
Common - SI TASKS: 2, 5, 6, 7, 11, 12, 18, 19, 35, 46, 49, 54, 55, 56, 64, 73, 75, 76, 77, 78, 79, 80, 96, 97, 98, 102, 124
Common for Respondents on both Q and SI TASKS: 5, 6, 7, 11, 12, 18, 19, 35, 46, 49, 55, 56, 64, 75, 76, 77, 98

3.50 - 4.49 PERFORMED WEEKLY
Graduate - Q TASKS: 1, 3, 4, 9, 10, 14, 15, 21, 27, 28, 29, 30, 41, 43, 44, 45, 65, 66, 120, 121, 124
Supervisor - Q TASKS: 1, 3, 9, 14, 15, 21, 29, 30, 41, 43, 44, 45, 65, 66, 120, 121
Common - Q TASKS: 1, 3, 9, 14, 15, 21, 29, 30, 41, 43, 44, 45, 65, 66, 120, 121, 124
Graduate - SI TASKS: 1, 3, 9, 14, 15, 21, 22, 23, 24, 25, 27, 28, 29, 30, 41, 42, 43, 44, 45, 48, 65, 66, 83, 95, 120
Supervisor - SI TASKS: 1, 9, 14, 15, 21, 41, 43, 45, 48, 65, 66, 72, 120, 121
Common - SI TASKS: 1, 9, 14, 15, 21, 41, 43, 45, 48, 65, 66, 120
Common for Respondents on both Q and SI TASKS: 1, 9, 14, 15, 21, 41, 43, 45, 65, 66, 120

4.50 - 5.00 PERFORMED DAILY
Graduate - Q TASKS: 22, 23, 24, 25, 42, 119
Supervisor - Q TASKS: 4, 10, 22, 23, 24, 25, 27, 28, 42, 119
Common - Q TASKS: 22, 23, 24, 25, 42, 119
Graduate - SI TASKS: 4, 10, 119
Supervisor - SI TASKS: 3, 4, 10, 22, 23, 24, 25, 27, 28, 29, 30, 42, 119
Common - SI TASKS: 4, 10, 119
Common for Respondents on both Q and SI TASKS: 119

APPENDIX C-2

TASKS RATED WITHIN SPECIFIED SUBCATEGORIES OF TRAINING ADEQUACY SCALE BY AEs

The numbers of tasks rated within specified training adequacy subcategories without regard to aircraft type or maintenance level are presented in this appendix. Table C-2 summarizes the number of tasks grouped according to the mean values by instrument (Q and SI) and by respondent (G and S). The number of tasks on which both graduates and supervisors agreed in their subcategory assignments is also shown. The final column lists the number of tasks on which all respondents agreed regardless of instrument used. The numbers in parentheses identify the mean values for those tasks rated in these subcategories.

TABLE C-2. NUMBER OF TASKS RATED WITHIN SPECIFIED TRAINING ADEQUACY SUBCATEGORIES BY GRADUATES (G) AND SUPERVISORS (S)

Subcategories	Q			SI			All Agree
	G	S	Agree	G	S	Agree	
1.00 - 1.49 Much More Emphasis	0	0	--	0	0	--	--
1.50 - 2.49 Increase Emphasis	23 (2.34)	20 (2.28)	13	26 (2.34)	12 (2.23)	8	6
2.50 - 3.49 Training Adequate	101 (2.78)	104 (2.92)	94	98 (2.75)	112 (2.91)	95	83
3.50 - 4.49 More Than Adequate	0	0	--	0	0	--	--
4.50 - 5.00 Reduce Training	0	0	--	0	0	--	--

Also shown in this appendix are lists of the numbers of the task statements rated by AEs. Response data are grouped by respondent, instrument, and tasks rated in common. The meanings ascribed to the five mean value subcategories are provided as subheadings.

1.00 - 1.49 Task Requires Much More Emphasis in "A" School
Graduate - Q TASKS: None
Supervisor - Q TASKS: None
Common - Q TASKS: None
Graduate - SI TASKS: None
Supervisor - SI TASKS: None
Common - SI TASKS: None
Common for Respondents on both Q and SI TASKS: None

1.50 - 2.49 Training Less Than Adequate for Task. Increase Emphasis in "A" School
Graduate - Q TASKS: 5, 6, 22, 23, 24, 25, 43, 48, 49, 72, 100, 101, 102, 107, 108, 109, 111, 112, 114, 119, 120, 122, 124
Supervisor - Q TASKS: 6, 7, 11, 13, 22, 23, 24, 25, 45, 48, 56, 72, 112, 114, 115, 119, 120, 122, 123, 124
Common - Q TASKS: 6, 22, 23, 24, 25, 48, 72, 112, 114, 119, 120, 122, 124
Graduate - SI TASKS: 5, 6, 11, 12, 22, 23, 24, 25, 26, 35, 45, 46, 47, 48, 49, 56, 61, 64, 67, 99, 100, 102, 115, 119, 122, 123
Supervisor - SI TASKS: 7, 13, 22, 23, 24, 25, 56, 57, 119, 120, 122, 123
Common - SI TASKS: 22, 23, 24, 25, 56, 119, 122, 123
Common for Respondents on both Q and SI TASKS: 22, 23, 24, 25, 119, 122

2.50 - 3.49 "A" School Training Adequate for Task
Graduate - Q TASKS: 1, 2, 3, 4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 44, 45, 46, 47, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 103, 104, 105, 106, 110, 113, 115, 116, 117, 118, 121, 123
Supervisor - Q TASKS: 1, 2, 3, 4, 5, 8, 9, 10, 12, 14, 15, 16, 17, 18, 19, 20, 21, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 46, 47, 49, 50, 51, 52, 53, 54, 55, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 103, 104, 105, 106, 107, 108, 109, 110, 111, 113, 116, 117, 118, 121
Common - Q TASKS: 1, 2, 3, 4, 8, 9, 10, 12, 14, 15, 16, 17, 18, 19, 20, 21, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 44, 46, 47, 50, 51, 52, 53, 54, 55, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 103, 104, 105, 106, 110, 113, 116, 117, 118, 121
Graduate - SI TASKS: 1, 2, 3, 4, 7, 8, 9, 10, 13, 14, 15, 16, 17, 18, 19, 20, 21, 27, 28, 29, 30, 31, 32, 33, 34, 36, 37, 38, 39, 40, 41, 42, 43, 44, 50, 51, 52, 53, 54, 55, 57, 58, 59, 60, 62, 63, 65, 66, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 101, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 120, 121, 124
Supervisor - SI TASKS: 1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 20, 21, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 121, 124

2.50 - 3.49 (continued)
Common - SI TASKS: 1, 2, 3, 4, 8, 9, 10, 14, 15, 16, 17, 18, 19, 20, 21, 27, 28, 29, 30, 31, 32, 33, 34, 36, 37, 38, 39, 40, 41, 42, 43, 44, 50, 51, 52, 53, 54, 55, 57, 58, 59, 60, 62, 63, 65, 66, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 101, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 116, 117, 118, 121, 124
Common for Respondents on both Q and SI TASKS: 1, 2, 3, 4, 8, 9, 10, 14, 15, 16, 17, 18, 19, 20, 21, 27, 28, 29, 30, 31, 32, 33, 34, 36, 37, 38, 39, 40, 41, 42, 44, 50, 51, 52, 53, 54, 55, 57, 58, 59, 60, 62, 63, 65, 66, 68, 69, 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 103, 104, 105, 106, 110, 113, 116, 117, 118, 121

3.50 - 4.49 Training More Than Adequate for Task. Reduce Emphasis in "A" School
Graduate - Q TASKS: None
Supervisor - Q TASKS: None
Common - Q TASKS: None
Graduate - SI TASKS: None
Supervisor - SI TASKS: None
Common for Respondents on both Q and SI TASKS: None

4.50 - 5.00 Greatly Reduce or Eliminate "A" School Training for This Task
Graduate - Q TASKS: None
Supervisor - Q TASKS: None
Common - Q TASKS: None
Graduate - SI TASKS: None
Supervisor - SI TASKS: None
Common - SI TASKS: None
Common for Respondents on both Q and SI TASKS: None

APPENDIX C-3

TASKS RATED WITHIN SPECIFIED SUBCATEGORIES OF PROFICIENCY SCALE BY AE SUPERVISORS

Proficiency ratings were obtained only from supervisors of AEs. Their response data are summarized in table C-3 without regard to aircraft type or maintenance level. The number of tasks on which supervisors agreed in their subcategory assignments is shown regardless of instrument used. The numbers in parentheses identify the mean values for those tasks rated in these subcategories.

TABLE C-3. NUMBER OF TASKS RATED WITHIN SPECIFIED PROFICIENCY SUBCATEGORIES

Subcategories	Q	SI	All Agree
1.00 - 1.49 Limited Performance Requires Direct Supervision	0	0	--
1.50 - 2.49 Performs w/Difficulty General Supervision	3 (2.39)	1 (2.47)	0
2.50 - 3.49 Partially Proficient Little Supervision	76 (3.12)	75 (3.09)	61
3.50 - 4.49 Performance Adequate w/o Supervision	45 (3.81)	42 (3.78)	30
4.50 - 5.00 Performs w/Ease Highly Proficient	0	6 (4.86)	--

Also presented in this appendix are lists of the tasks rated which are grouped into the same mean value subcategories. The responses are grouped by respondent, instrument, and common ratings. The meanings associated with the mean values are provided as category subheadings.

1.00 - 1.49 Extremely Limited in Performing Task, Cannot Do Without Direct Supervision
Graduate - Q TASKS: None
Supervisor - Q TASKS: None
Common - Q TASKS: None

1.50 - 2.49 Performs Task With Difficulty, Requires General Supervision
Supervisor - Q TASKS: 115, 118, 123
Supervisor - SI TASKS: 122
Common for Respondents of both Q and SI TASKS: None

2.50 - 3.49 Partially Proficient, Requires Little Supervision
Supervisor - Q TASKS: 3, 5, 6, 7, 11, 12, 13, 20, 21, 22, 23, 24, 25, 26, 31, 32, 33, 34 36, 44, 45, 46, 47, 48, 49, 50, 51, 53, 54, 55, 56, 58, 59, 60, 61, 62, 63, 64, 67, 68, 69, 70, 71, 72, 73, 74, 75, 78, 79, 82, 83, 84, 88, 89, 90, 92, 94, 97, 99, 100, 101, 102, 104, 107, 108, 110, 111, 112, 113, 114, 116, 117, 119, 120, 122, 124
Supervisor - SI TASKS: 5, 6, 7, 8, 11, 12, 13, 20, 21, 22, 23, 24, 25, 26, 28, 29, 30, 32, 33, 34, 36, 38, 39, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 58, 59, 60, 61, 62, 63, 64, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 79, 90, 92, 99, 100, 101, 102, 103, 104, 105, 107, 109, 111, 112, 113, 114, 115, 119, 120, 121, 123, 124
Common for Respondents on both Q and SI TASKS: 5, 6, 7, 11, 12, 13, 20, 21, 22, 23, 24, 25, 26, 32, 33, 34, 36, 44, 45, 46, 47, 48, 49, 50, 51, 53, 54, 55, 56, 58, 59, 60, 61, 62, 63, 64, 67, 68, 69, 70, 71, 72, 73, 74, 75, 79, 90, 92, 99, 100, 101, 102, 104, 107, 111, 112, 113, 114, 119, 120, 124

3.50 - 4.49 Performance Is Adequate Without Supervision
Supervisor - Q TASKS: 1, 2, 4, 8, 9, 10, 14, 15, 16, 17, 18, 19, 27, 28, 29, 30, 35, 37, 38, 39, 40, 41, 42, 43, 52, 57, 65, 66, 76, 77, 80, 81, 85, 86, 87, 91, 93, 95, 96, 98, 103, 105, 106, 109, 121
Supervisor - SI TASKS: 3, 4, 9, 10, 14, 15, 16, 17, 18, 19, 27, 31, 35, 37, 40, 41, 42, 43, 65, 66, 77, 78, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 91, 93, 94, 95, 96, 97, 98, 106, 108, 110
Common for Respondents on both Q and SI TASKS: 4, 9, 10, 14, 15, 16, 17, 18, 19, 27, 35, 37, 40, 41, 42, 43, 65, 66, 77, 80, 81, 85, 86, 87, 91, 93, 95, 96, 98, 106
4.50 - 5.00 Performs Task With Ease, Highly Proficient
Supervisor - Q TASKS: None
Supervisor - SI TASKS: None
Common for Respondents on both Q and SI TASKS: None

APPENDIX D-1

TASKS RATED WITHIN SPECIFIED SUBCATEGORIES OF FREQUENCY SCALE BY BTs

Response data for the frequency of task scale are presented in this appendix without regard to watch station qualification or steaming conditions. Table D-1 shows the number of tasks rated within the five subcategories. The response selections are grouped according to instrument (Q and SI) and type of respondent (G and S). Also shown are the number of tasks on which graduates and supervisors agreed in their subcategory assignments. The final column lists the number of tasks on which all respondents agreed regardless of instrument used. The numbers in parentheses identify the mean values for those tasks rated in these subcategories.

TABLE D-1. NUMBER OF TASKS RATED WITHIN SPECIFIED FREQUENCY SUBCATEGORIES BY GRADUATES (G) AND SUPERVISORS (S)

Subcategories	Q			SI			All Agree
	G	S	Agree	G	S	Agree	
1.00 - 1.49 Never Performed	8 (1.34)	4 (1.33)	2	11 (1.28)	19 (1.25)	10	1
1.50 - 2.49 Seldom Performed	54 (2.11)	43 (2.07)	35	62 (2.06)	55 (1.93)	50	24
2.50 - 3.49 Performed Monthly	39 (2.95)	50 (2.94)	31	37 (2.87)	31 (2.96)	27	17
3.50 - 4.49 Performed Weekly	21 (3.90)	22 (3.95)	14	14 (3.74)	16 (3.87)	11	7
4.50 - 5.00 Performed Daily	2 (4.59)	5 (4.70)	2	0	3 (4.56)	—	—

The appendix also contains the identity of the task statements grouped in the same five mean value categories. The task numbers are presented by instrument and by respondent.

1.00 - 1.49 NEVER PERFORMED
Graduate - Q TASKS: 4, 69, 70, 77, 78, 104, 118, 119
Supervisor - Q TASKS: 4, 27, 104, 123
Common - Q TASKS: 4, 104
Graduate - SI TASKS: 4, 38, 69, 70, 73, 77, 97, 104, 118, 119, 123
Supervisor - SI TASKS: 4, 34, 38, 45, 69, 70, 73, 77, 78, 85, 97, 98, 104, 111, 117, 118, 119, 123, 124
Common - SI TASKS: 38, 69, 70, 73, 77, 97, 104, 118, 119, 123
Common for all Respondents on both Q and SI TASKS: 104

1.50 - 2.49 SELDOM PERFORMED OR ONLY IN EMERGENCIES
Graduate - Q TASKS: 3, 6, 7, 8, 9, 15, 17, 19, 24, 27, 33, 34, 37, 38, 39, 42, 45, 48, 58, 59, 61, 64, 71, 73, 76, 83, 84, 85, 86, 87, 89, 90, 91, 92, 93, 97, 98, 101, 103, 105, 106, 107, 108, 109, 110, 111, 112, 116, 117, 120, 121, 122, 123, 124
Supervisor - Q TASKS: 3, 7, 8, 29, 33, 34, 37, 38, 39, 40, 42, 45, 64, 69, 70, 71, 73, 76, 77, 78, 33, 84, 85, 86, 89, 90, 91, 92, 97, 98, 101, 102, 105, 107, 108, 109, 110, 111, 117, 118, 119, 121, 124
Common - Q TASKS: 3, 7, 8, 33, 34, 35, 37, 38, 39, 42, 45, 64, 71, 73, 76, 83, 84, 85, 86, 89, 90, 91, 92, 97, 98, 101, 105, 107, 108, 109, 110, 111, 117, 121, 124
Graduate - SI TASKS: 3, 6, 7, 8, 9, 14, 15, 17, 19, 23, 24, 27, 28, 29, 31, 33, 34, 37, 39, 40, 42, 45, 47, 48, 49, 50, 58, 59, 61, 64, 71, 76, 78, 79, 83, 84, 85, 86, 87, 89, 90, 91, 92, 93, 98, 101, 103, 105, 106, 107, 108, 109, 110, 111, 112, 115, 116, 117, 120, 121, 122, 124
Supervisor - SI TASKS: 3, 6, 7, 8, 9, 15, 17, 19, 23, 24, 27, 28, 29, 31, 33, 37, 39, 40, 42, 50, 58, 59, 61, 64, 66, 71, 76, 79, 83, 84, 86, 87, 88, 89, 90, 91, 92, 93, 100, 101, 102, 103, 105, 105, 107, 108, 109, 110, 112, 114, 115, 116, 120, 121, 122
Common - SI TASKS: 3, 7, 8, 9, 15, 17, 19, 23, 24, 27, 28, 29, 31, 33, 37, 39, 40, 42, 50, 58, 59, 61, 64, 71, 76, 78, 79, 83, 84, 86, 87, 89, 90, 91, 92, 93, 101, 103, 105, 106, 107, 108, 109, 110, 112, 115, 116, 120, 121, 122
Common for all Respondents on both Q and SI TASKS: 3, 7, 8, 33, 37, 39, 42, 64, 71, 76, 83, 84, 86, 89, 90, 91, 92, 101, 105, 107, 108, 109, 110, 121

2.50 - 3.49 PERFORMED MONTHLY
Graduate - Q TASKS: 5, 10, 11, 13, 14, 16, 18, 20, 22, 23, 25, 28, 29, 30, 31, 32, 40, 43, 47, 50, 56, 57, 62, 63, 66, 68, 74, 75, 79, 82, 88, 94, 95, 99, 100, 102, 113, 114, 115
Supervisor - Q TASKS: 2, 6, 9, 14, 15, 16, 17, 18, 19, 20, 22, 23, 24, 25, 26, 28, 30, 31, 47, 48, 50, 56, 57, 58, 59, 61, 62, 63, 66, 74, 75, 79, 82, 87, 88, 93, 94, 95, 96, 99, 100, 103, 106, 112, 113, 114, 115, 116, 120, 122
Common - Q TASKS: 14, 16, 18, 20, 22, 23, 25, 28, 30, 31, 32, 47, 50, 56, 57, 62, 63, 66, 74, 75, 79, 82, 88, 94, 95, 96, 99, 100, 113, 114, 115
Graduate - SI TASKS: 2, 5, 10, 11, 12, 13, 16, 18, 20, 22, 25, 26, 30, 32, 36, 43, 56, 57, 62, 63, 65, 66, 67, 68, 74, 75, 81, 82, 88, 94, 95, 96, 99, 100, 102, 113, 114
Supervisor - SI TASKS: 2, 10, 11, 13, 14, 16, 18, 20, 22, 25, 26, 30, 36, 47, 48, 49, 56, 57, 62, 63, 65, 67, 74, 75, 81, 82, 94, 95, 96, 99, 113
Common - SI TASKS: 2, 10, 11, 13, 16, 18, 20, 22, 25, 26, 30, 36, 56, 57, 62, 63, 65, 67, 74, 76, 82, 94, 95, 96, 99, 113, 114
Common for all Respondents on both Q and SI TASKS: 16, 18, 20, 22, 25, 30, 56, 57, 62, 63, 82, 94, 95, 96, 99, 113, 114

3.50 - 4.49 PERFORMED WEEKLY
Graduate - Q TASKS: 1, 2, 12, 21, 26, 35, 36, 41, 44, 46, 49, 52, 54, 55, 60, 65, 67, 72, 80, 81, 96
Supervisor - Q TASKS: 1, 5, 10, 11, 12, 13, 21, 32, 35, 36, 41, 43, 44, 46, 49, 52, 65, 67, 68, 72, 80, 81
Common - Q TASKS: 1, 12, 21, 36, 41, 44, 46, 49, 52, 65, 67, 72, 80, 81
Graduate - SI TASKS: 1, 21, 35, 41, 44, 46, 51, 52, 53, 54, 55, 60, 72, 80
Supervisor - SI TASKS: 1, 5, 12, 21, 32, 35, 41, 43, 44, 46, 51, 52, 55, 60, 68, 80
Common - SI TASKS: 1, 21, 35, 41, 44, 46, 51, 52, 55, 60, 80
Common for all Respondents on both Q and SI TASKS: 1, 21, 41, 44, 46, 52, 80

4.50 - 5.00 PERFORMED DAILY
Graduate - Q TASKS: 51, 53
Supervisor - Q TASKS: 51, 53, 54, 55, 60
Common - Q TASKS: 51, 53
Graduate - SI TASKS: None
Supervisor - SI TASKS: 53, 54, 72
Common - SI TASKS: None
Common for all Respondents on both Q and SI TASKS: None

APPENDIX D-2

TASKS RATED WITHIN SPECIFIED SUBCATEGORIES OF TRAINING ADEQUACY SCALE BY BTs

In this appendix response data for the adequacy of BT "A" School training is listed without regard to watch qualification or steaming conditions. Table D-2 presents the number of tasks assigned to five subcategories of ratings. These are grouped according to instrument (Q and SI) and respondent (G and S). The number of tasks which the respondents commonly selected to be in a given category is shown also as a function of the instrument type and finally without regard to instrument. The numbers in parentheses identify the mean values for those tasks rated in these subcategories.

TABLE D-2. NUMBER OF TASKS RATED WITHIN SPECIFIED TRAINING ADEQUACY SUBCATEGORIES BY GRADUATES (G) AND SUPERVISORS (S)

Subcategories	Q			SI			All Agree
	G	S	Agree	G	S	Agree	
1.00 - 1.49 Much More Emphasis	0	0	--	0	3 (1.31)	--	--
1.50 - 2.49 Increase Emphasis	77 (2.21)	19 (2.34)	16	87 (2.23)	60 (2.08)	55	14
2.50 - 3.49 Training Adequate	46 (2.68)	103 (2.90)	46	36 (2.72)	60 (2.70)	32	26
3.50 - 4.49 More Than Adequate	1 (3.65)	2 (3.64)	1	1 (3.54)	1 (4.30)	1	1
4.50 - 5.00 Reduce Training	0	0	--	0	0	--	--

The task numbers which identify the statements rated are also provided in this appendix. They are grouped according to respondent and instrument within the five mean value subcategories.

1.00 - 1.49 Task Requires Much More Emphasis in "A" School
Graduate - Q TASKS: None
Supervisor - Q TASKS: None
Common - Q TASKS: None
Graduate - SI TASKS: None
Supervisor - SI TASKS: 59, 118, 123
Common - SI TASKS: None
Common for all Respondents on both Q and SI TASKS: None

1.50 - 2.49 Training Less Than Adequate for Task. Increase Emphasis in "A" School
Graduate - Q TASKS: 2, 3, 7, 8, 11, 12, 13, 15, 16, 17, 18, 19, 23, 25, 26, 27, 28, 29, 31, 33, 34, 37, 38, 39, 45, 47, 49, 50, 51, 58, 59, 60, 61, 62, 63, 64, 65, 66, 69, 70, 71, 73, 74, 75, 76, 77, 78, 79, 85, 86, 87, 89, 90, 91, 92, 93, 96, 97, 98, 101, 102, 103, 104, 105, 107, 108, 109, 111, 115, 116, 117, 118, 119, 121, 122, 123, 124
Supervisor - Q TASKS: 23, 27, 37, 38, 45, 58, 59, 74, 76, 77, 78, 89, 97, 98, 116, 118, 119, 121, 124
Common - Q TASKS: 23, 37, 38, 45, 58, 59, 76, 77, 78, 89, 97, 98, 116, 118, 119, 121
Graduate - SI TASKS: 2, 3, 5, 6, 7, 8, 9, 11, 13, 15, 16, 17, 19, 20, 23, 25, 26, 27, 28, 29, 30, 31, 32, 34, 37, 38, 39, 40, 45, 47, 48, 49, 50, 51, 58, 59, 60, 61, 62, 63, 64, 65, 66, 68, 69, 70, 71, 73, 74, 75, 76, 77, 78, 79, 82, 85, 86, 87, 88, 89, 90, 91, 92, 93, 96, 97, 98, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124
Supervisor - SI TASKS: 1, 7, 8, 9, 15, 17, 19, 23, 28, 29, 34, 37, 38, 39, 40, 45, 47, 49, 50, 58, 61, 62, 63, 64, 65, 66, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 97, 98, 101, 102, 111, 115, 116, 117, 119, 120, 121, 122, 124
Common - SI TASKS: 7, 8, 9, 15, 17, 19, 23, 28, 29, 34, 37, 38, 39, 40, 45, 47, 49, 50, 58, 61, 62, 63, 64, 65, 66, 69, 70, 71, 73, 74, 75, 76, 77, 78, 82, 85, 86, 87, 88, 89, 90, 91, 92, 97, 98, 101, 102, 115, 116, 117, 119, 120, 121, 122, 124
Common for all Respondents on both Q and SI TASKS: 23, 37, 38, 45, 58, 76, 77, 78, 89, 97, 98, 116, 119, 121

2.50 - 3.49 "A" School Training Adequate for Task
Graduate - Q TASKS: 1, 5, 6, 9, 10, 14, 20, 21, 22, 24, 30, 32, 35, 36, 40, 41, 42, 43, 44, 46, 48, 52, 53, 54, 55, 56, 57, 67, 68, 72, 80, 81, 82, 83, 84, 88, 94, 95, 99, 100, 106, 110, 112, 113, 114, 120
Supervisor - Q TASKS: 1, 2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 24, 25, 26, 28, 29, 30, 31, 32, 33, 34, 35, 36, 39, 40, 41, 42, 43, 44, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 90, 91, 92, 93, 94, 95, 96, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 117, 120, 122, 123
Common - Q TASKS: 1, 5, 6, 9, 10, 14, 20, 21, 22, 24, 30, 32, 35, 36, 40, 41, 42, 43, 44, 48, 52, 53, 54, 55, 56, 57, 67, 68, 72, 79, 80, 81, 82, 83, 84, 88, 94, 95, 99, 100, 106, 110, 112, 113, 114, 120
Graduate - SI TASKS: 1, 10, 12, 14, 18, 21, 22, 24, 33, 35, 36, 41, 42, 43, 44, 46, 52, 53, 54, 55, 56, 57, 67, 72, 80, 81, 83, 84, 94, 95, 99, 110, 111, 112, 113, 114
Supervisor - SI TASKS: 2, 3, 5, 6, 10, 11, 12, 13, 14, 16, 18, 20, 21, 22, 24, 25, 26, 27, 30, 31, 32, 33, 35, 36, 41, 42, 43, 44, 46, 48, 51, 52, 53, 54, 55, 56, 57, 60, 67, 68, 79, 80, 81, 93, 94, 95, 96, 99, 100, 103, 104, 105, 106, 107, 108, 109, 110, 112, 113, 114
Common - SI TASKS: 10, 12, 14, 18, 21, 22, 24, 33, 35, 36, 41, 42, 43, 44, 46, 51, 52, 53, 54, 55, 56, 57, 63, 67, 80, 94, 95, 99, 110, 112, 113, 114
Common for all Respondents on both Q and SI TASKS: 10, 14, 21, 22, 24, 35, 36, 41, 42, 43, 44, 52, 53, 54, 55, 56, 57, 67, 80, 94, 95, 99, 110, 112, 113, 114

3.50 - 4.49 Training More Than Adequate for Task. Reduce Emphasis in "A" School
Graduate - Q TASKS: 4
Supervisor - Q TASKS: 4, 46
Common - Q TASKS: 4
Graduate - SI TASKS: 4
Supervisor - SI TASKS: 4
Common - SI TASKS: 4
Common for all Respondents on both Q and SI TASKS: 4

4.50 - 5.00 Greatly Reduce or Eliminate "A" School Training for This Task
Graduate - Q TASKS: None
Supervisor - Q TASKS: None
Common - Q TASKS: None
Graduate - SI TASKS: None
Supervisor - SI TASKS: None
Common - SI TASKS: None
Common for all Respondents on both Q and SI TASKS: None

111

APPENDIX D-3

TASKS RATED WITHIN SPECIFIED SUBCATEGORIES OF PROFICIENCY SCALE BY BT SUPERVISORS

Proficiency ratings were obtained only from supervisors of BTs. The response data are summarized in table D-3 without regard to watch qualification or steaming conditions. The number of tasks assigned to each subcategory is shown first in terms of the instrument used and, secondly, without regard to the instrument. The numbers in parentheses identify the mean values for those tasks rated in these subcategories.

TABLE D-3. NUMBER OF TASKS RATED WITHIN SPECIFIED PROFICIENCY SUBCATEGORIES

Subcategories	Q	SI	All Agree
1.00 - 1.49 Limited Performance Requires Direct Supervision	0	4 (1.06)	--
1.50 - 2.49 Performs w/Difficulty General Supervision	1 (2.21)	53 (2.14)	1
2.50 - 3.49 Partially Proficient Little Supervision	95 (3.04)	67 (2.87)	39
3.50 - 4.49 Performance Adequate w/o Supervision	28 (3.72)	0	--
4.50 - 5.00 Performs w/Ease Highly Proficient	0	0	--

This appendix also provides lists of task statement numbers which were assigned to each subcategory in terms of the instrument (Q and SI) and respondent (G and S). Lists are provided which indicate areas having common ratings.

1.00 - 1.49 Extremely Limited in Performing Task, Cannot Do Without Direct Supervision
Supervisor - Q TASKS: None
Supervisor - SI TASKS: 59, 73, 77, 123
Common for all Respondents on both Q and SI TASKS: None

1.50 - 2.49 Performs Task With Difficulty, Requires General Supervision
Supervisor - Q TASKS: 38
Supervisor - SI TASKS: 1, 2, 3, 4, 6, 7, 8, 9, 15, 16, 17, 19, 23, 28, 34, 37, 38, 39, 40, 45, 47, 50, 58, 61, 63, 64, 66, 69, 70, 72, 74, 76, 78, 82, 84, 85, 86, 87, 88, 89, 90, 91, 92, 97, 98, 101, 111, 116, 117, 118, 119, 121, 124
Common for all Respondents on both Q and SI TASKS: 38

2.50 - 3.49 Partially Proficient, Requires Little Supervision
Supervisor - Q TASKS: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 23, 24, 25, 26, 27, 28, 29, 30, 31, 33, 34, 37, 39, 40, 42, 45, 48, 50, 57, 58, 59, 61, 62, 63, 64, 66, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 96, 97, 98, 99, 101, 102, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124
Supervisor - SI TASKS: 5, 10, 11, 12, 13, 14, 18, 20, 21, 22, 24, 25, 26, 27, 29, 30, 31, 32, 33, 35, 36, 41, 42, 43, 44, 46, 48, 49, 51, 52, 53, 54, 55, 56, 57, 60, 62, 65, 67, 68, 71, 75, 79, 80, 81, 83, 93, 94, 95, 96, 99, 100, 102, 103, 104, 105, 106, 107, 108, 109, 110, 112, 113, 114, 115, 120, 122
Common for all Respondents on both Q and SI TASKS: 10, 11, 12, 13, 14, 18, 20, 21, 24, 25, 26, 27, 29, 30, 31, 33, 42, 48, 57, 62, 83, 93, 94, 96, 99, 102, 104, 105, 106, 107, 108, 109, 110, 112, 113, 114, 115, 120, 122

3.50 - 4.49 Performance Is Adequate Without Supervision
Supervisor - Q TASKS: 6, 12, 22, 32, 35, 36, 41, 43, 44, 46, 47, 49, 51, 52, 53, 54, 55, 56, 60, 65, 67, 68, 79, 80, 81, 95, 100, 103
Supervisor - SI TASKS: None
Common for all Respondents on both Q and SI TASKS: None

4.50 - 5.00 Performs Task With Ease, Highly Proficient
Supervisor - Q TASKS: None
Supervisor - SI TASKS: None
Common for all Respondents on both Q and SI TASKS: None

APPENDIX E

SCALE INDEPENDENCE: COMPARATIVE DATA

This appendix contains a list of correlations which are relevant to the issue of scale independence. The numerics identify the specific scales of tables 4 and 8 involved in each correlation. Correlations between ratings of training adequacy and proficiency are not included.

SCALE INDEPENDENCE: COMPARATIVE DATA

CORRELATION	AE		BT
r1,2	.04	vs.	.48
r1,4	-.16	vs.	.51
r1,5	.03	vs.	.68
r2,3	.02	vs.	.48
r3,4	-.22	vs.	.55
r3,5	-.07	vs.	.71
r6,2	-.03	vs.	.44
r6,4	-.27	vs.	.48
r6,5	-.08	vs.	.66
r7,1	-.04	vs.	.58
r7,3	-.11	vs.	.55
r8,2	.01	vs.	.41
r8,4	-.22	vs.	.45
r8,5	-.04	vs.	.63
r9,1	-.17	vs.	.54
r9,3	-.25	vs.	.60
r10,1	.32	vs.	.69
r10,3	.23	vs.	.73
r6,7	-.11	vs.	.50
r6,9	-.28	vs.	.55
r6,10	.23	vs.	.72
r7,8	-.07	vs.	.49
r8,9	-.24	vs.	.59
r8,10	.29	vs.	.75

APPENDIX F

OVERALL SCALE MEANS AND STANDARD DEVIATIONS

This appendix contains overall scale means and standard deviations. Each overall mean was calculated from the mean scale values of the 124 tasks for each rating. They are arranged to allow comparisons of scale values by instrument for each rate as well as between rate comparisons.

TABLE F-1. OVERALL SCALE MEANS AND STANDARD DEVIATIONS

	Graduate				Supervisor					
	Frequency		Adequacy		Frequency		Adequacy		Proficiency	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
AE										
Q	2.717	.961	2.703	.254	2.586	1.015	2.815	.321	3.357	.437
SI	2.632	.993	2.666	.244	2.724	1.015	2.841	.264	3.407	.520
BT										
Q	2.671	.811	2.398	.315	2.839	.850	2.842	.326	3.186	.379
SI	2.426	.701	2.389	.295	2.411	.912	2.377	.485	2.498	.516

DISTRIBUTION LIST

Air Force

Headquarters, Air Training Command (XPTD, Dr. D. E. Meyer)
Headquarters, Air Training Command (XPTIA, Mr. Goldman)
Air Force Human Resources Laboratory, Brooks Air Force Base
Air Force Human Resources Laboratory (Library), Lowry Air Force Base
Air Force Office of Scientific Research/AR (Dr. A. R. Fregly)

Army

Commandant, TRADOC (Technical Library)
ARI (Dr. Ralph R. Canter, 316C; Dr. Edgar Johnson; Mr. James Baker)
ARI Field Unit - Fort Leavenworth

Coast Guard

Commandant, U.S. Coast Guard Headquarters (G-P-1/62; G-RT/81)
U.S. Coast Guard Training Center, Governors Island (Ms. J. Chippendale)

Marine Corps

CMC (OT)
CGMCDEC (Mr. Greenup)
Director, Marine Corps Institute
CO MARCORCOMMELECSCOL (ITS) (Mr. D. E. Mueller)

Navy

OASN (R&D) (Dr. S. Koslov) (MRA&L) (Dr. B. D. Rostker)
CNO (OP-102X, M. Malehorn; OP-987P10, Dr. R. Smith; OP-987, Mr. H. Stone)
NAVCOMPT (NCD-7)
ONR (458, 455)
CNM (MAT-08T2, Mr. A. L. Rubinstein)
CNET (01, 00A, N-5 (6 copies), 015 (5 copies))
CNAVRES (02)
COMNAVSEASYSCOM (03, 047C1, 047C12)
COMNAVAIRSYSCOM (03, 340F)
CO NAVMEDRSCHDEVCOM (CAPT H. J. Connery)
CNET SUPPORT (00, 00A (2 copies))
CNTECHTRA (016, L. DeVeau (5 copies), R. Coolidge, 0161, Dr. Kerr)
CNATRA (Dr. F. Schufletowski)
COMTRALANT
COMTRALANT (Educational Advisor)

DISTRIBUTION LIST (continued)

Navy (continued)

COMTRAPAC (2 copies)
CO NAVPERSRANDCEN (Library (4 copies))
NAVPERSRANDCEN Liaison (OIH)
Superintendent NAVPGSCOL (2124)
Superintendent Naval Academy Annapolis (Chairman, Behavioral Science Dept.)
CO NAVEDTRAPRODEVCEN (AH3)
CO NAVEDTRASUPPCEN NORVA
CO NAVEDTRASUPPCENPAC (5 copies)
CO NAVAEROMEDRSCHLAB (Chief Aviation Psych. Div.)
CO FLTCOMDIRSYSTRACENPAC
CO NAMTRAGRU
CISO, NTTC Corry Station
CO NAVTRAEQUIPCEN (N-424 (2 copies))
Center for Naval Analyses (2 copies)
U.S. Naval Institute (CDR Bowler)
OIC NODAC (2)
CO TRITRAFAC (2 copies)
CO NAVSUBTRACENPAC (2 copies)
CO FLEASWTRACENPAC
CO FLETRACEN SDIEGO
CISO, SSC (Dr. Flesch)
Executive Director NAVINSTPRODEVDET
CO NAVTECHTRACEN Corry Station (Cryptologic Training Department)
OIC CTS for PTEP, Guided Missile School, Dam Neck (2 copies)

Other DOD

Military Assistant for Human Resources, OAD (E&LS), Room 3D129, Pentagon
(CDR Paul Chatelier)
OASD (MRA&L) (LT COL Grossel)
Director, Human Resources Office, Defense Advanced Research Projects Agency
(Dr. Dexter Fletcher)
Institute for Defense Analyses (Dr. Jesse Orlansky)

Information Exchanges

DDC (12 copies)
DLSIE (Mr. James Dowling)
Executive Editor, Psychological Abstracts, American Psychological Association
ERIC Processing and Reference Facility, Bethesda, MD (2 copies)